

1

FACTOR 3

THE DIMENSIONALITY OF NATIONS PROJECT

RESEARCH REPORT

FACTOR 2

FACTOR 1

DEPARTMENT OF POLITICAL SCIENCE

UNIVERSITY OF HAWAII

DISTRIBUTION STATEMENT A

Approved for public release;
AFG. Distribution Unlimited

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

RECEIVED
JUN 21 1971
C

54

DISCLAIMER NOTICE

THIS DOCUMENT IS THE BEST
QUALITY AVAILABLE.

COPY FURNISHED CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

The Dimensionality of Nations Project

University of Hawaii

RESEARCH REPORT NO. 15

Investigations into Alternative Techniques for
Developing Empirical Taxonomies:
The Results of Two Plasmodes

Warren R. Phillips

October 1968

Prepared in Connection with the National Science Foundation,
Grant No. GS-1230, and the Advanced Research Projects Agency,
ARPA Order No. 1063, and Monitored by the Office of Naval
Research, Contract #N00014-67-A-0387-0003 ✓

This document has been approved for public release and sale; its
distribution is unlimited and reproduction in whole or in part is
permitted for any purpose of the United States Government

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	1
1. INTRODUCTION	2
2. PLASMODE	3
2.1 Factor Groups	4
2.2 Hierarchical Clustering Scheme	16
2.3 Conclusions from Plasnodes	25
3. SUBSTANTIVE EXAMPLE	31

~~Unclassified~~

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

The Dimensionality of Nations Project
University of Hawaii
2500 Campus Road, Honolulu, Hawaii 96822

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

Investigations into Alternative Techniques for Developing Empirical Taxonomies:
The Results of Two Plasmodes

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report No. 15

5. AUTHOR(S) (First name, middle initial, last name)

Warren R. Phillips

6. REPORT DATE

October 1968

7a. TOTAL NO. OF PAGES

51

7b. NO. OF REFS

17

8a. CONTRACT OR GRANT NO.

N00014-67-A-0387-0003

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Research Report No. 15

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited and reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES

2500 Campus Road
Honolulu, Hawaii 96822

12. SPONSORING MILITARY ACTIVITY

Advanced Research Projects Agency
Washington, D.C.

13. ABSTRACT

This is a research report on the applicability of grouping techniques to the data of international relations. While it is still too early to make strong recommendations for the "best" grouping method, the results of this paper suggest some important characteristics of current techniques which should be considered in selecting a grouping technique.

Several goals guided this research: First, to demonstrate the characteristics of current clustering techniques by employing experiments with known outcomes--plasmodes; secondly, to suggest a measure of similarity that is advantageous in grouping experiments when correlations are of less meaning; and thirdly, to employ the techniques investigated in the plasmode to a substantive problem from international relations.

Data were collected on the mileage between a number of cities in the United States. The cities were arranged into two matrixes: one of 22 cities, which divided into obvious groups; the other of 60 cities, with no discernible groups. The distance matrix was then rescaled to a similarity matrix varying for 0.0 to 1.0, where 1.0 is the closest. The principal components of this matrix were computed and the similarities matrix was also hierarchically decomposed employing an algorithm supplied by S. C. Johnson. The groups delineated by these techniques were compared geographically on maps of the United States.

The research suggests that direct factor analysis techniques seem more appealing than the hierarchical clustering schemes for describing the structure of the spaces defined by the plasmodes. Hierarchical clustering schemes are useful when the researcher wishes to break up a dense cluster of entities, however.

DD FORM 1473 (PAGE 1)

5/11 6101-807-6801

~~Unclassified~~

Security Classification

Unclassified
Security Classification

Security Classification

DD FORM 1473 (BACK)
(PAGE 2)

Security Classification

List of Tables and Figures

		<u>Page</u>
Table 1	List of Cities in Plasmodes 1 and 2	5
Table 2	Orthogonal Rotation (22 cities)	6
Table 3	Orthogonal Rotation (60 cities)	7
Figure I	Orthogonal Grouping (22 cities)	9
Figure II	Orthogonal Grouping (60 cities)	10
Table 4	Oblique Rotation (22 cities)	12
Table 5	Oblique Rotation (60 cities)	13
Figure III	Oblique Grouping (22 cities)	15
Figure IV	Oblique Grouping (66 cities)	16
Figure V	Dendrogram for 22 cities (Diameter Method)	21
Figure VI	Dendrogram for 22 cities (Connectedness Method)	22
Figure VII	HCS Grouping of 22 cities (Diameter Method)	23
Figure VIII	HCS Grouping of 22 cities (Connectedness Method)	24
Figure IX	Dendrogram for 60 cities (Diameter Method)	26
Figure X	Dendrogram for 60 cities (Connectedness Method)	27
Figure XI	HCS Grouping of 60 cities (Diameter Method)	28
Figure XII	HCS Grouping of 60 cities (Connectedness Method)	29
Figure XIII	Flow Chart of Dyadic Conflict Analysis	33
Table 6	Orthogonal Rotation of Conflict Dyads	34
Figure XIV	Orthogonal Grouping of Conflict Dyads	35
Table 7	Oblique Rotation of Conflict Dyads	37
Figure XV	Oblique Grouping of Conflict Dyads	38
Figure XVII	Dendrogram for Conflict Dyads (Diameter Method)	40
Figure XVIII	HCS Grouping of Conflict Dyads (Diameter Method)	42
Figure XIX	HCS Grouping of Conflict Dyads (Connectedness Method)	43

ABSTRACT:

This is a research report on the applicability of grouping techniques to the data of international relations. While it is still too early to make strong recommendations for the "best" grouping method, the results of this paper suggest some important characteristics of current techniques which should be considered in selecting a grouping technique.

Several goals guided this research: First, to demonstrate the characteristics of current clustering techniques by employing experiments with known outcomes---plasmodes; secondly, to suggest a measure of similarity that is advantageous in grouping experiments when correlations are of less meaning; and thirdly, to employ the techniques investigated in the plasmodes to a substantive problem from international relations.

Data were collected on the mileage between a number of cities in the United States. The cities were arranged into two matrixes: one of 22 cities, which divided into obvious groups; the other of 60 cities, with no discernible groups. The distance matrix was then rescaled to a similarity matrix varying for 0.0 to 1.0, where 1.0 is the closest. The principal components of this matrix were computed and the similarities matrix was also hierarchically decomposed employing an algorithm supplied by S. C. Johnson. The groups delineated by these techniques were compared geographically on maps of the United States.

The research suggests that direct factor analysis techniques seem more appealing than the hierarchical clustering schemes for describing the structure of the spaces defined by the plasmodes. Hierarchical clustering schemes are useful when the researcher wishes to break up a dense cluster of entities, however.

1. INTRODUCTION

Grouping nations, objects, individuals, or cases by types is a basic step in describing phenomenon and building science. The virtue of typing is that it enables parsimonious descriptions of objects and facilitates reliable predictions about them based upon their group identity. Classification is the process of ordering cases into groups that best represent certain empirically measured relations of contiguity, similarity, or both. While this process is certainly not new to the social sciences, it has been more readily associated with biological taxonomies. In taxonomic work in biology--where typal distinctions have been built in a non-systematic (non-quantitative) fashion, it is becoming increasingly clear that a more systematic basis is required for scientifically meaningful classification (Sokal and Sneath, 1963).

In political science voices are being raised for the use of systematic grouping techniques (Russett 1967, Rummel 1969, Brams, 1966). The problem with the prevailing types in political science is that the rationale underlying the categorization is often not explicit. It is not clear whether our "types" really divide different kinds of variance. If we are to deal in types a clear and empirical basis for the distinctions must be made.

Several techniques for grouping entities have been suggested recently.¹ I have chosen two techniques which seem to be most fruitful for use in international relations. In order to investigate the applicability of these techniques, a special type of "experiment" has been designed. This type of experiment is one in which the outcome is known before the experimenter begins his analysis and is termed a plasmode. Plasmodes are used when an analyst is interested in discerning the characteristics of his analytical tools and not the characteristics of the entities under study. Cattell suggests the use

of plasmodes to investigate the properties of a specific method (Cattell, 1967). No comments:

The usefulness of a plasmode resides first in helping us evaluate the power of various experimental and analytical procedures to show that the data in any experiment does not fit a particular model. For, before applying the investigatory procedures to be evaluated, we know already, by virtue of our "back stage" information, what actual structures should be revealed. From seeing the degree to which they are revealed, we are able to evaluate the relative soundness of the given methods and inferential procedures. In all ordinary investigations, as distinct from investigations with a plasmode, the scientist leans on the belief that his experimental methods and statistical analyses are capable of revealing whether the data has a structure which fits some particular model, and on this assumption he proceeds to "substantive" discoveries in the given domain. In working with a plasmode, precisely the converse approach is made, since the technique not the substantive area is investigated. He knows exactly what the "substantive" structural answer must be, and he is setting out to discover things about the method... (pp. 304)

2. PLASMODE

In order to keep the structure of a plasmode in its simplest form, a two dimensional problem was chosen for this investigation. When we deal with only two dimensions, we can plot entities' positions and then visually search for clusters. For this purpose a common road map mileage chart of several U. S. cities seems ideal. The mileage chart becomes a matrix of distances between all of the cities identified in the matrix. The matrix is square (cities across cities) and symmetric (the distance from Chicago to Detroit is the same as from Detroit to Chicago). The grouping of cities ought to be quite easily recognizable. There may be more than one way to define groups, depending on the size of the groups we wish to identify and on the general criteria we happen to apply. Any of these choices can be tested, however, against a visual inspection of the groups on a map to ascertain their logical "meaning."

Two plasmodes were employed. The first contains 22 cities, chosen to group into rather distinct sets. The second matrix contains 60 cities, which are spread out across the map. This second matrix allows us to ascertain the tendency of the techniques to describe groups which do not seem to be intuitively meaningful. Table I contains the list of cities for both plasmodes.

2.1 Factor Groups

We will present two methods of grouping nations on their distances. The first method is a direct factor analysis of the distance matrix (Rummel, 1969, Section 22.2). The (mileage) distances are first scaled to lie between 0+1.00, where 1.00 is the smallest distance. This transforms the distance matrix to a similarity matrix, which is then factor analyzed (principal components) as though it were a correlation matrix. The resulting factors define cities whose pattern of distances from other cities are interdependent--similar in profile. Cities with high loadings on the same factor are similarly located in space.

The factors can be rotated to orthogonal and oblique simple structure to ascertain the clearest definition of groups. Using factor analysis in this fashion brings out distinct groups: there would be no ambiguity as to the number of types or the membership of each type.²

In employing factor analysis, the principle axis technique was applied to the similarities (transformed mileage) matrix. Rotation to orthogonal solution was performed on three factors in both experiments. The choice of this cutoff will be discussed later. Tables II and III present the orthogonal rotations of the 22 city and 60 city plasmodes.

In the 22 city example, the factors define three clusters. The first cluster is centered around Albany, New York, and defines a group of eastern and mid-western cities. The second factor, centered around Colorado Springs and

Table 1. List of Cities in Plasmode 1 and 2

<u>Plasmode I</u>	<u>Plasmode II</u>	<u>Plasmode II(cont.)</u>
1 Albany, N. Y.	1 Albuquerque, N. M.	31 Los Angeles, Calif.
2 Ashville, Tenn.	2 Atlanta, Ga.	32 Louisville, Ky.
3 Atlanta, Ga.	3 Baltimore, Md.	33 Memphis, Tenn.
4 Atlantic City, N. J.	4 Birmingham, Ala.	34 Milwaukee, Wisc.
5 Baltimore, Md.	5 Bismarck, N.D.	35 Minneapolis, St. Paul, Minn.
6 Binghampton, N.Y.	6 Boston, Mass.	36 Nashville, Tenn.
7 Birmingham, Ala.	7 Buffalo, N.Y.	37 New Orleans, La.
8 Boston, Mass.	8 Cheyenne, Wyo.	38 New York, N. Y.
9 Buffalo, N.Y.	9 Chicago, Ill.	39 Omaha, Neb.
10 Charlotte, N.C.	10 Cincinnati, Ohio	40 Peoria, Ill.
11 Chattanooga, Tenn.	11 Cleveland, Ohio	41 Philadelphia, Pa.
12 Chicago, Ill.	12 Columbus, Ohio	42 Phoenix, Ariz.
13 Cincinnati, Ohio	13 Dallas, Tex.	43 Pierre, S.D.
14 Cleveland, Ohio	14 Denver, Colo.	44 Pittsburgh, Pa.
15 Colorado Sps., Colo.	15 Des Moines, Iowa	45 Portland, Ore.
16 Columbus, Ohio	16 Detroit, Mich.	46 Raleigh, S.C.
17 Dallas, Texas	17 Dubuque, Iowa	47 St. Louis, Mo.
18 Davenport, Iowa	18 Evansville, Ind.	48 Salt Lake, Utah
19 Denver, Colo.	19 Ft. Wayne, Ind.	49 San Francisco, Cal.
20 Detroit, Mich.	20 Grand Canyon, N.M.	50 Seattle, Wash.
21 El Paso, Tex.	21 Glacier Nat. Park, Mont.	51 Spokane, Wash.
	22 Helena, Mont.	52 Springfield, Ill.
	23 Houston, Tex.	53 Springfield, Mo.
	24 Indianapolis, Ind.	54 Tampa, Fla.
	25 Jackson, Miss.	55 Toledo, Ohio
	26 Jacksonville, Fla.	56 Topeka, Kansas
	27 Jefferson City, Mo.	57 Tulsa, Okla.
	28 Kansas City, Mo.	58 Washington, D.C.
	29 Lansing, Mich.	59 Wichita, Kan.
	30 Little Rock, Ark.	60 Yellowstone, Wyo.

Table 2. Orthogonal Rotation (22 cities)

<u>City</u>		1	Factor 2	3
Albany	1	9203	0284	2150
Asheville	2	4976	1939	7964
Atlanta, Ga.	3	3838	2437	8672
Atlantic City	4	8420	0240	3736
Baltimore	5	8253	0927	4366
Bingham	6	9137	0772	2673
Birmingham	7	3501	3416	8240
Boston	8	8750	-0600	2307
Buffalo	9	9059	2006	2502
Charlotte	10	5321	1293	7733
Chatanooga	11	4410	3034	8095
Chicago	12	6719	4831	3836
Cincinnati	13	6808	3631	5396
Cleveland	14	8346	2916	3659
Colorado Sp.	15	1593	9253	1181
Columbus	16	7514	3411	4701
Dallas	17	1317	7084	5262
Davenport	18	5953	5701	3525
Denver	19	2012	9304	0743
Detroit	20	7968	3353	3571
El Paso	21	-0745	8274	2609
Erie	22	8769	2409	3045

Table 3. Orthogonal Rotation (60 Cities)

Variable No.	Name	Communality 3 Factors	Factor		
			1	2	3
1	Albuquerque	824	209	645	-604
2	Atlanta	870	684	087	-629
3	Baltimore	871	903	066	-227
4	Birmingham	894	646	134	-678
5	Bismarck	801	505	720	-167
6	Boston	767	871	-001	-095
7	Buffalo	880	907	156	-179
8	Cheyenne	841	375	750	-370
9	Chicago	912	814	366	-340
10	Cincinnati	921	839	234	-402
11	Cleveland	924	904	216	-245
12	Columbus	922	870	214	-345
13	Dallas	889	379	387	-771
14	Denver	835	345	722	-442
15	Des Moines	868	661	511	-559
16	Detroit	911	886	256	-245
17	Dubuque	876	747	446	-344
18	Evansville	910	746	270	-530
19	Ft. Wayne	927	855	289	-336
20	Grand Canyon	805	053	733	-515
21	Glacier Nat.	847	250	886	-006
22	Helena	870	249	895	-083
23	Houston	841	355	294	-793
24	Indiana	928	820	285	-417
25	Jackson	900	535	188	-760
26	Jacksonville	779	624	-023	-624
27	Jefferson City	891	643	407	-599
28	Kansas City	884	579	484	-561
29	Lausling	899	864	286	-266
30	Little Rock	908	538	300	-727
31	Los Angeles	725	-077	721	-446
32	Louisville	914	795	232	-479
33	Memphis	915	608	255	-692
34	Milwaukee	881	801	391	-295
35	Minn.-St. Paul	924	677	544	-265
36	Nashville	903	714	193	-597
37	New Orleans	849	485	151	-769
38	New York	821	890	024	-169
39	Omaha	858	595	566	-429
40	Peoria	899	753	387	-428
41	Philadelphia	850	900	046	-191
42	Phoenix	769	033	669	-566
43	Pierre	818	507	699	-269
44	Pittsburgh	913	909	161	-247
45	Portland	799	025	893	-008

Variable		Communality		Factor	
No.	Name	3 Factors	1	2	3
46	Raleigh	813	805	015	-406
47	St. Louis	909	702	354	-540
48	Salt Lake City	851	186	853	-296
49	San Francisco	700	-078	793	-239
50	Seattle	806	066	893	056
51	Spokane	856	164	910	-019
52	Springfield, I.	902	731	373	-478
53	Springfield, II.	898	581	400	-633
54	Tampa	737	563	-054	-645
55	Toledo	930	890	261	-263
56	Topeka	874	549	510	-559
57	Tulsa	884	484	436	-677
58	Washington	865	892	066	-253
59	Wichita	873	474	527	-617
60	Yellowstone	885	275	879	-190

Denver, is a group of Rocky Mountain and South Western cities, while the third factor is a southern dimension centered around Atlanta, Georgia.

In the 60 city case, the cutoff criteria for rotation was three factors. The fourth factor explained only 2.7 percent of the variance with no loadings above .5. The first cluster, by far the strongest, centered around Pittsburgh, Cleveland and Baltimore. This factor is centered in the East and groups cities both West and South. The second factor is centered in the Northwest at Spokane and covers the West. The third factor loosely centers around New Orleans and Houston. This factor has no high loadings on it. Most of the cities are included within the other two factors in the southern states both to the east and west.

The information presented in the factor tables can also be conveyed using a map of the United States. The observer can more easily check the structuring of the plasmode experiments with his intuitive "feel" for the grouping of cities based upon mileage distances.

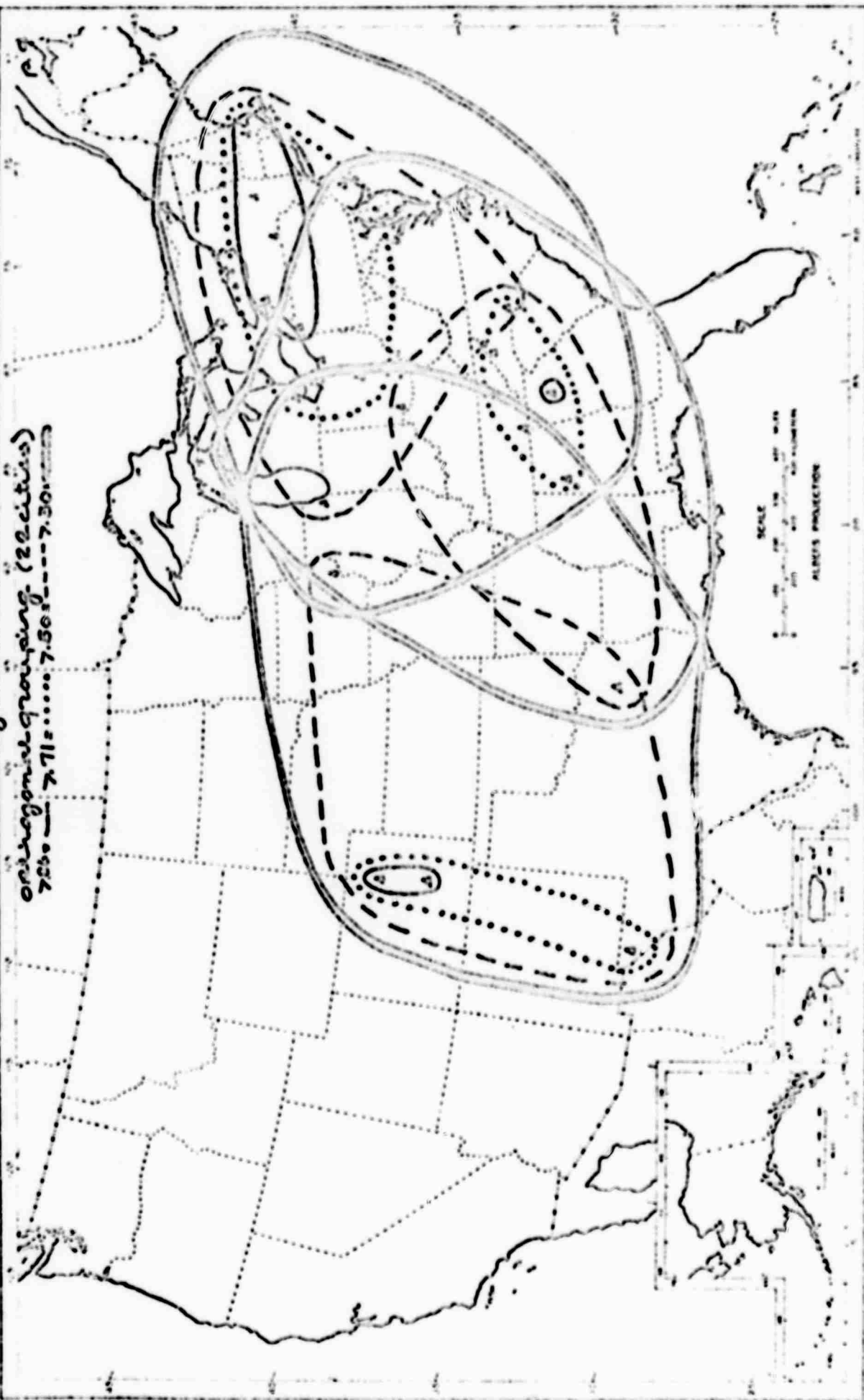
The groupings of cities are readily discernible in the map presentation. The contours or concentric circles on the map refer to factor loading criteria. Thus, the inner, thin circle clusters all cities with factor loadings above .86 on the same factor. The other circles refer to loading criteria of .71, .50, and .30 respectively. Turning to the groupings, a mid-western group might have been expected. These mid-western cities were included in the other groups, but they did not group into a separate cluster. In both plasmodes these cities would have grouped if a fourth factor had been employed. The percentage of variance accounted for in a fourth factor in each case was very low. The loadings were below .500 and the factors would not have been rotated under normal criteria of the strength of loadings or a SKEEL test (Cattell, 1966). They were not employed in the plasmode for the same reasons.

UNITED STATES

Figure I

No. 110

Orthogonal grouping (22 cities)
7550 — 7712 — 7568 — 7301



COAST AND GEOD. SURV.
WASHINGTON, D.C.
1960

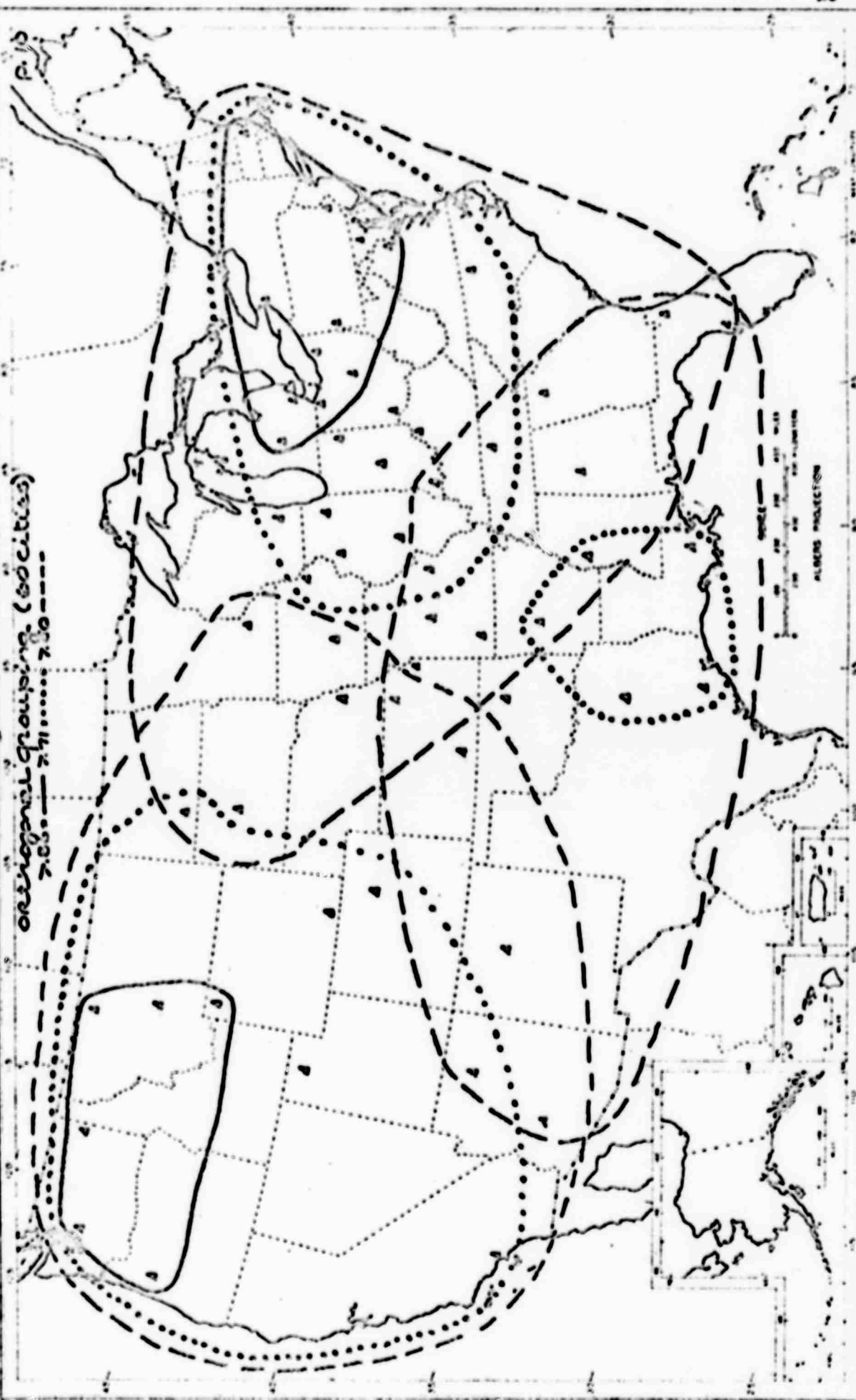
Figure I, No. 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200

No. 110

UNITED STATES

Figure II

Orthogonal grouping (societies)



GOODE BASE MAP SERIES
 UNITED STATES
 1:250,000
 1900

Prepared by Henry H. Rogers
 under the direction of the
 U.S. GEOLOGICAL SURVEY

Both factor structures were rotated to an oblique, bi-quartain solution. Oblique rotation was derived to ascertain the relationships between the group factors in this analysis. If the factors are highly intercorrelated, each group is overlapping the other groups and the factors are not equally defining a cluster of dyads. The primary pattern matrix has been interpreted in the 22 city experiments. Two sharp factors appeared with a third factor showing some moderate loadings. The three factors are quite similar to the orthogonal solution except that the first factor has six cities loading above .86, while in the orthogonal case there were only five. Albany and Binghamton are at the center of the first cluster. The second factor centers around Denver and Colorado Springs, while the third factor is a southern factor loosely centered at Atlanta.

The factor correlation matrix shows that factors 1 and 3 have a moderate correlation. Factor two is slightly less related to either factor 1 or

	1	2	3
1	_____		
2	.2679	_____	
3	.4322	.3210	_____

3. Thus the groups in the East and the South seem more associated with each other than either is with the western group.

In contrast to the mild changes associated with the shift in rotation on the 22 city matrix, oblique factor analysis of the 60 city matrix produced striking changes from the orthogonal analysis. In this placement there is only one well defined cluster. All five of the cities with loadings above .86--Bismark, Glacier, Helena, Pierre, and Yellowstone--are equally central to the cluster defined by the factor. Only eight of the 60 cities have a loading of less than .50 on this factor. It would appear that the one factor accounts for most of the spacial variation in cities.

Table 4. Oblique Rotation (22 Cities)

<u>City</u>		<u>1</u>	<u>2</u>	<u>3</u>
Albany	1	9807	-0986	-0342
Asheville	2	4035	0386	6973
Atlanta	3	2619	0932	7997
Atlantic City	4	8683	-1173	1624
Baltimore	5	8343	-0521	2252
Binghamton	6	9607	-0539	0173
Birmingham	7	2236	2060	7469
Boston	8	9352	-1891	0098
Buffalo	9	9448	0794	-0184
Charlotte	10	4508	-0299	6722
Chatanooga	11	3298	1581	7109
Chicago	12	6407	3830	1508
Cincinnati	13	6347	2336	3375
Cleveland	14	8394	1665	1144
Colorado Sp.	15	0808	9410	-0608
Columbus	16	7261	2127	2446
Dallas	17	0009	6578	4271
Davenport	18	5539	4872	1253
Denver	19	1340	9481	-1216
Detroit	20	7954	2178	1089
El Paso	21	-1934	8426	1778
Erie	22	9004	1172	0430

Table 5. Oblique Rotation (60 Cities)

<u>City</u>		<u>1</u>	Factor <u>2</u>	<u>3</u>
Albuquerque	1	6470	5698	2993
Atlanta	2	4655	0852	3358
Baltimore	3	5754	-3337	-0902
Birmingham	4	4823	1612	3785
Bismark	5	8794	1206	-2097
Boston	6	5020	-4363	-1811
Buffalo	7	6509	-3339	-1607
Cheyenne	8	8291	3516	0059
Chicago	9	7686	-0803	-0464
Cincinnati	10	6759	-1044	0403
Cleveland	11	6984	-2605	-1170
Columbus	12	6771	-1710	-0151
Dallas	13	5350	4845	4729
Denver	14	7885	4080	0884
Des Moines	15	7984	1189	0255
Detroit	16	7206	-2338	-1231
Dubuque	17	7959	-0061	-0454
Evansville	18	6512	0542	1746
Ft. Wayne	19	7293	-1381	-0404
Grand Canyon	20	6288	6328	2410
Glacier Nat.	21	8676	2214	-3255
Helena	22	8745	2803	-2576
Houston	23	4449	4756	5264
Indiana	24	7064	-0626	0447
Jackson	25	4625	3055	4719
Jacksonville	26	3414	0717	3812
Jefferson City	27	7033	1902	1934
Kansas City	28	7292	2600	1930
Lansing	29	7323	-1942	-1060
Little Rock	30	5557	3257	4086
Los Angeles	31	5439	6545	2221
Louisville	32	6478	-0257	1242
Memphis	33	5597	2427	3685
Milwaukee	34	7815	-0941	-0902
Minn.-St. Paul	35	8350	0186	-1236
Nashville	36	5700	0683	2669
New Orleans	37	4032	3249	5055
New York	38	5334	-3848	-1260
Omaha	39	8056	1914	0449
Peoria	40	7507	0261	0464
Philadelphia	41	5579	-3657	-1162
Phoenix	42	5650	6543	3119
Pierre	43	8636	1830	-1119
Pittsburgh	44	6559	-2849	-1012
Portland	45	7438	5563	-2559
Raleigh	46	4766	-1713	1174
St. Louis	47	6943	1207	1728
Salt Lake City	48	8040	4508	-0325

<u>City</u>		<u>1</u>	Factor <u>2</u>	<u>3</u>
San Francisco	49	6065	5397	0120
Seattle	50	7678	2870	-3271
Spokane	51	8377	2639	-3280
Springfield, I.	52	7264	0684	1026
Springfield, M.	53	6624	2755	2811
Tampa	54	2809	1090	4283
Toledo	55	7271	-2209	-1090
Topeka	56	7333	2866	1930
Tulsa	57	6359	3778	3412
Washington	58	5695	-3090	-0631
Wichita	59	6953	3741	2670
Yellowstone	60	8769	3347	-1638

In addition to the major factor, two minor factors appear. The first is centered around the Grand Canyon and the second around New Orleans and Houston. These two minor factors are bipolar in that the loadings define a group of cities that are positively loaded on each factor and another group that are negatively loaded on the factors. The implication of bipolar factors will become clearer when we turn to the geographic displays.

In interpreting the oblique rotation in the 60 city plasmode, the reference structure matrix was employed rather than the more customary primary pattern matrix. Two criteria were influential on this decision. The primary pattern matrix had 29 loadings above .86 and all but three loadings above .50. The interpretation of grouping procedures would have been very difficult with the primary pattern. The factor correlation matrices also provided information for choosing between the two rotations.

The primary pattern correlation matrix:

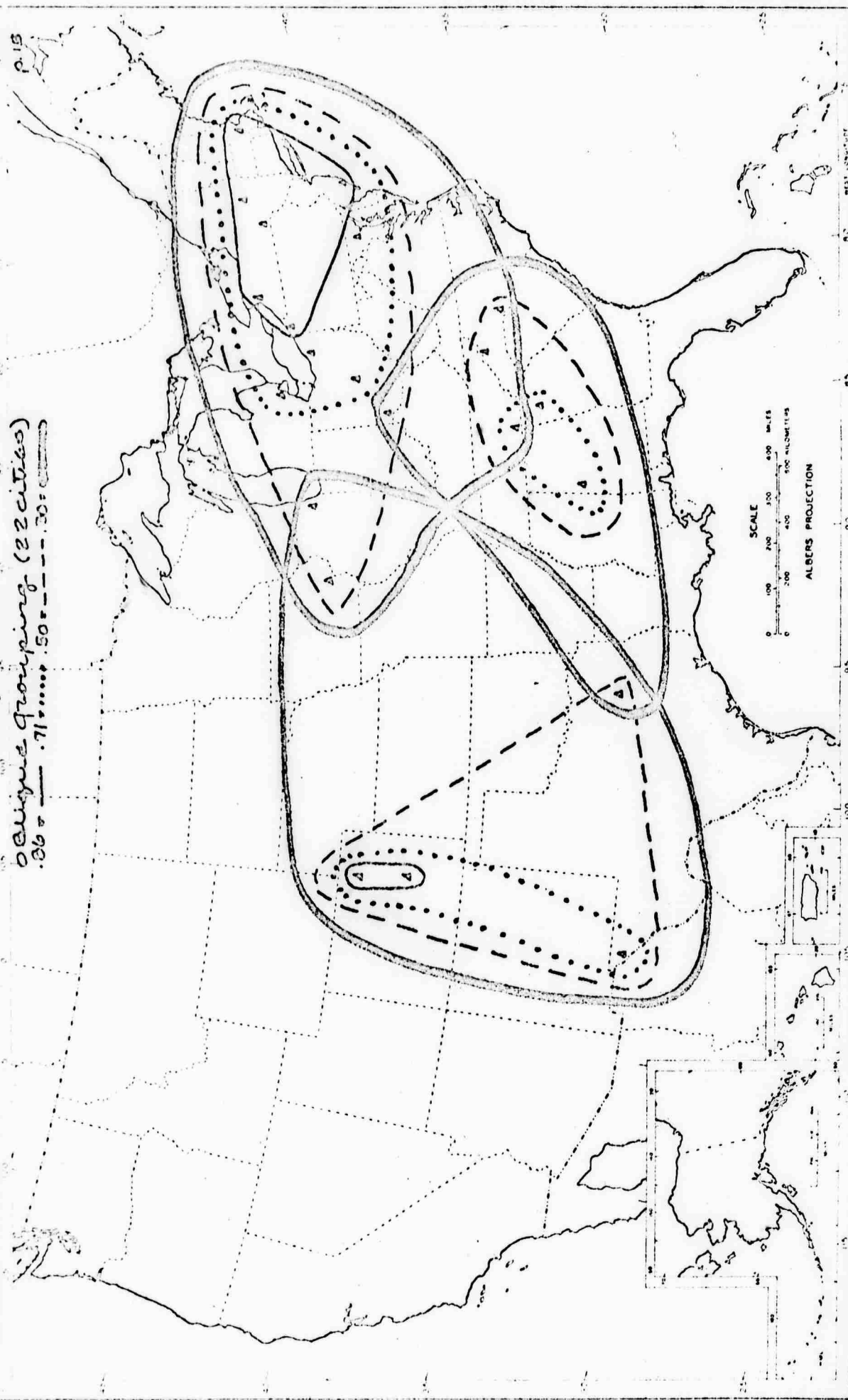
Factor	1	2	3
1	_____		
2	-.44	_____	
3	.58	-.77	_____

The reference structure correlation matrix:

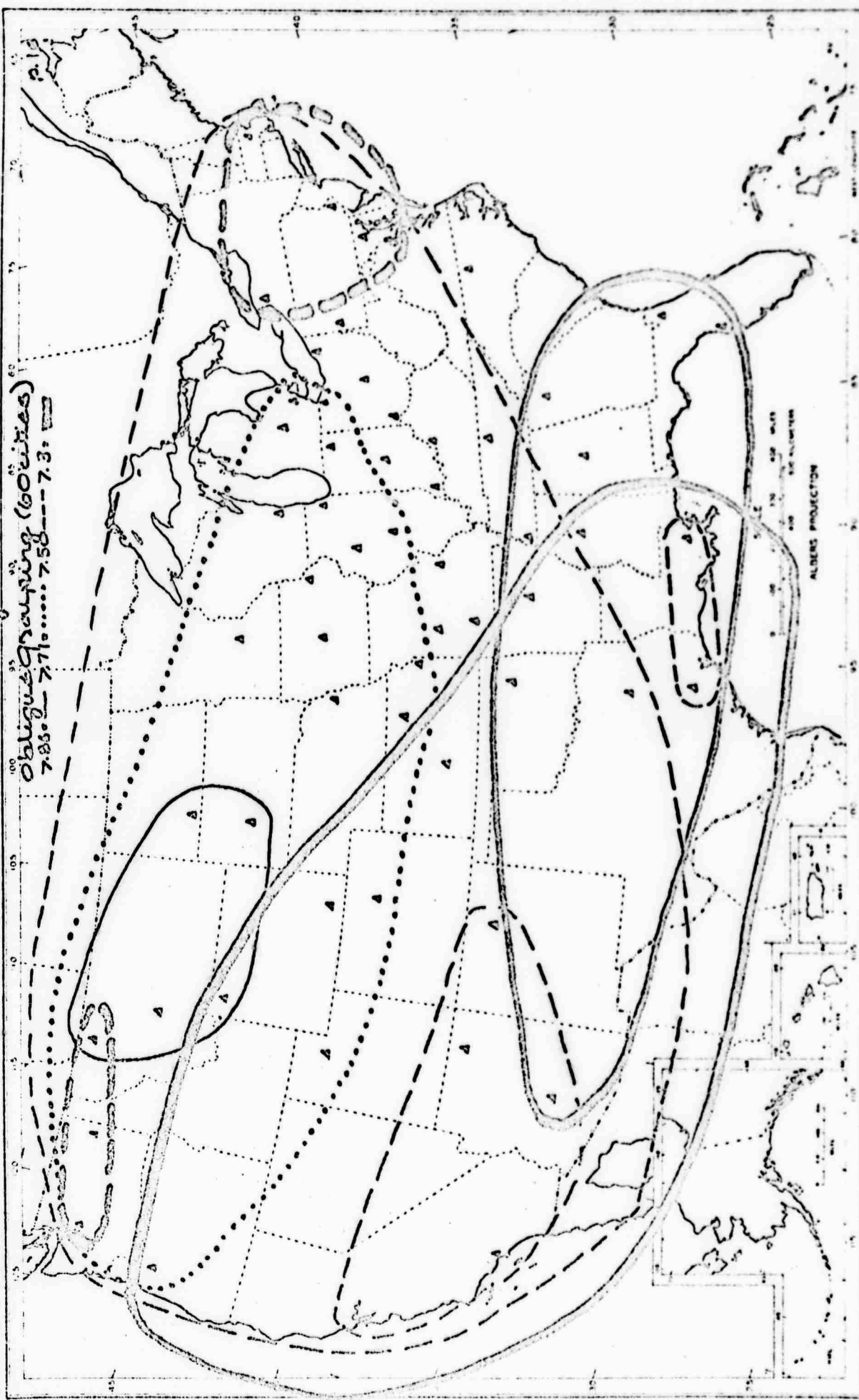
Factor			
1	_____		
2	-.00	_____	
3	-.41	.70	_____

By discussing the reference structure matrix we are able to reduce the correlation between factors 1 and 2 and in general reduce all correlations between factors. Because of the very high loadings for the primary factors and the high correlations between the primary factors, the reference structure matrix was selected as giving a clearer simple structure.

Oblique Grouping (22 cities)
 .36 - .71 50 - 100 - 150 - 200 - 250 - 300 - 350 - 400 - 450 - 500 - 550 - 600 - 650 - 700 - 750 - 800 - 850 - 900 - 950 - 1000



GOODE BASE MAP SERIES
 PREPARED BY HENRY M. LEPPARD
 1900 for the University of Chicago



The geographic presentation of the two plasnodes demonstrates the grouping of cities in both plasnodes. In the plasnode with 22 cities there is little change in the geographic location of groups from the orthogonal case. The most notable change is the shrinkage in size of the two smaller groups; both the Southern and Western clusters come more sharply in focus.

In the 60 city plasnode more striking changes occur, as we should expect, given the differences which appeared in the interpretation of the factor matrices. Now, there is one large grouping centered in the Northwest; cities group from the East and South as the loading criteria is relaxed from .86 to .71 and finally to .50. The two weak factors are centered in the Southwest and the South. In grouping cities according to loadings, two small groups appear which have opposite loadings on each of the minor factors. One of these groups--designated by dashed lines is in the East and represents the bipolar loadings on the southwestern factor. The other small group is in the far Northwest and is the reverse of the Southern cluster.

While all the groups are interpretable, oblique solutions do not seem to be quite as meaningful when there are no clear cut clusters of cities as in the 60 city plasnode. The analysis tends to produce groups which are too large to interpret easily. This is due to the fact that a large group of cities tend to have the same pattern of distances across all cities and therefore cluster together even though they are quite distant from each other.

2.2 Hierarchical Clustering Scheme

Another method for grouping entities is based upon the distance between all points in space taken one at a time rather than the patterns of variation in distances across all entities associated with factor analysis. Factor analysis looks at a matrix of cities as a series of vectors.

Thus, factor analysis begins with a square symmetric matrix-- cities by cities, in this case--of similarities. The resulting factors define cities whose pattern of distances from other cities are interdependent-- similar in profile. Cities with high loadings on the same factor display similar patterns of distances down their respective column vectors and are thus similarly located in space. In contrast, this new method builds a hierarchical grouping of cities and produces a taxonomic tree or dendrogram with cities that are closest in distance on the bottom of each branch. This technique is closely associated with biological sciences; the methodology is the same as that employed in building taxonomies of reptiles, mammals or insects (Sokal and Sneath, 1963).

Several techniques by which a dendrogram can be built have been published. The most recent of them is by S. C. Johnson (1967). It subsumes Ward's (1963) grouping technique and that employed by Sokal and Sneath (1963). The technique operates on either a distance or similarities matrix. It groups objects on the basis of their distances from each other. Thus the closest two points in space will be grouped together first. The next closest entities are grouped secondly, and so forth, until all entities are in a single group. The key to the method is replacing two (or more) objects (cities) with a single entity (cluster) that defines the distance between the newly formed cluster and all other cities or clusters. When two cities form a group, there are two alternative distances between the group and the other entities: the original distances between both of the cities that joined to form the group and the other entities not in the group. Johnson's program allows the choice of either the maximum distance (the greater of the two distances) or the minimum distance. He terms these alternative choices the Diameter and Connectedness methods respectively. The resulting taxonomies are invariant

under monotone transformations of the similarity data. Thus, we need have only rank order confidence in our data.

The diameter method attempts to minimize the diameter of the groups. Thus the diameter method builds groups by adding a city to a group if the maximum distance between the city and group members is smaller than that between the group and other cities not in the group. Every time a group is formed between two cities x and y there is a choice as to which distance we should use to represent the distance from the group $/x, y/$ and other cities such as z . The diameter method makes the choice of the maximum distance

$$d([x,y],z) = \max [d(x,z), d(y,z)].$$

where d = distance

$[x,y]$ = a cluster of cities x,y

z = another city

The connectedness method, however, adds a city to a group if the minimum distance between the city and any member of the group is smaller than that between the group and other cities not in the group. In its choice of distances between the group and other cities the criteria employed is the minimum distance

$$d(/x,y/,z) = \min [d(x,z), d(y,z)]$$

where d = distance

$/x,y/$ = a cluster of cities x,y

z = another city

The connectedness method tends to build chains of entities that are 'sausage-shaped', with the length of the sausage chain minimized.

Using the same plasmodes analyzed with factor analysis I have employed Johnson's Hierarchical Clustering Scheme (HCS) on the original distance matrices. I will display the results in two forms. First the dendrograms will be described and then the geometric interpretations of the dendrograms will be presented on maps of the United States.

The form chosen to present the dendrogram needs clarification. The program print-out includes a dendrogram which has a level for each different distance (similarity) in the matrix. These levels or branches in the dendrogram are too numerous for interpretation and comparison with previous results. In order to make the hierarchical clustering analyses comparable with the factor analysis results, I calculated the mean (\bar{x}) number of cities that grouped at each of the four levels of clustering in the factor results (.86, .71, .50, and .30). The mean size of each of the factor groups was used as a criteria for defining groups in hierarchical clustering. For instance, the mean size of the tightest cluster groups (.86) in the factor analysis of the 22 city plasmode was three cities. The first grouping of at least three cities in the dendrogram was chosen as the first cluster. These levels are identified as K levels in the figures. At each K level additional groups that have formed since the last level are identified regardless of size.

The dendrograms for both HCS techniques on the 22 city plasmode are reported in figures V and VI.

The diameter method has grouped all cities before the last level (.22). The connectedness method does not pull in El Paso and Dallas until the .27 level. Figures VII and VIII present the geometric displays.

The diameter method groups cities into 4 meaningful groups; the East, Midwest, South, and West. The connectedness method does not make as clear a grouping of cities, however. The major grouping is in the East-Midwestern area,

Figure V. Dendogram for 22 Cities

Connectedness Method

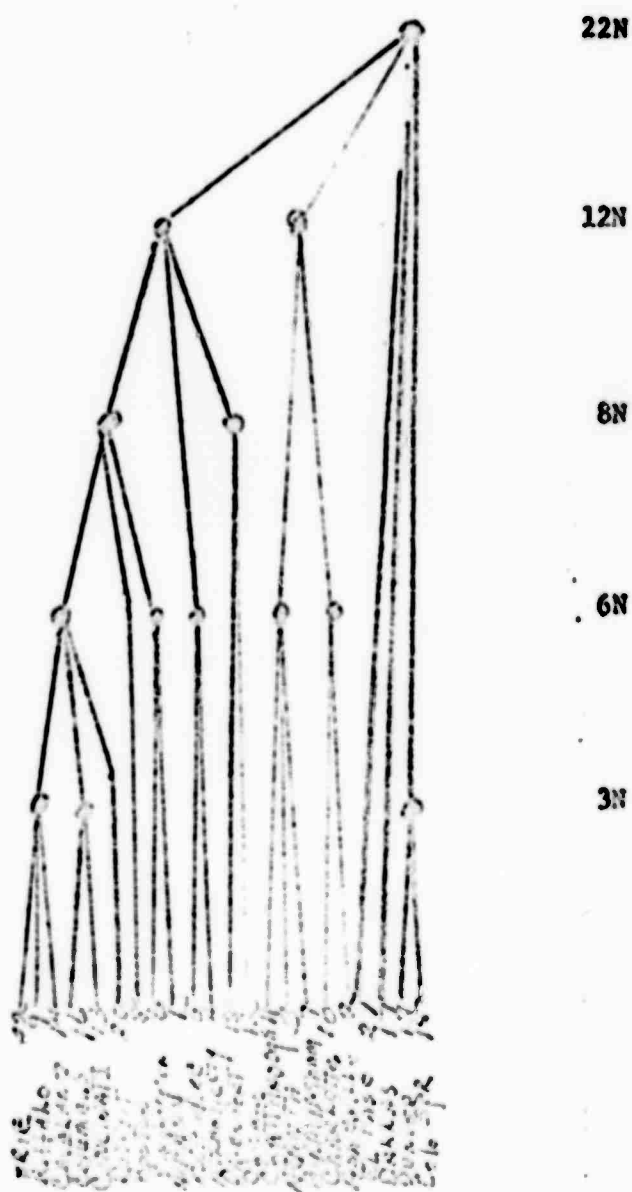
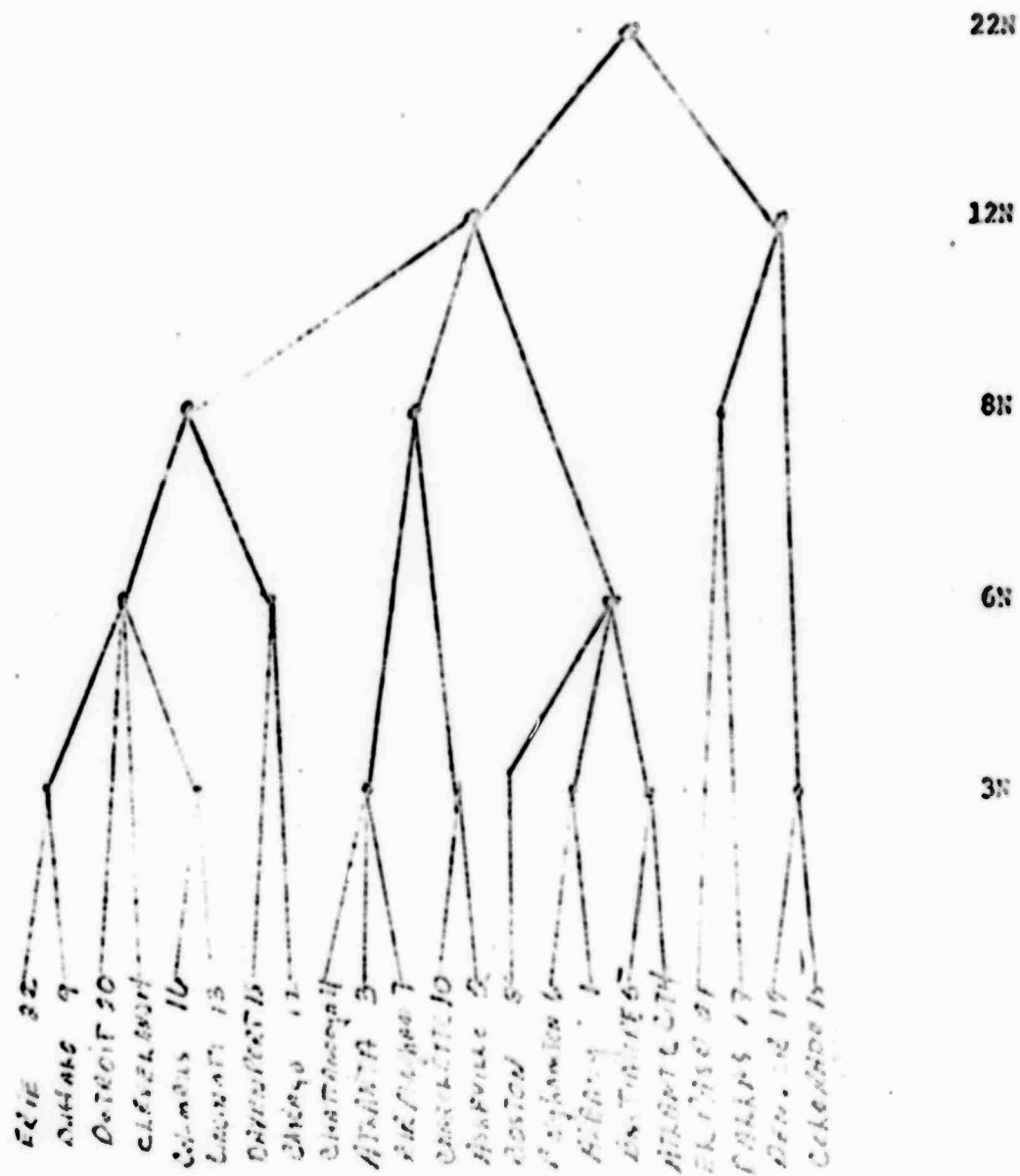
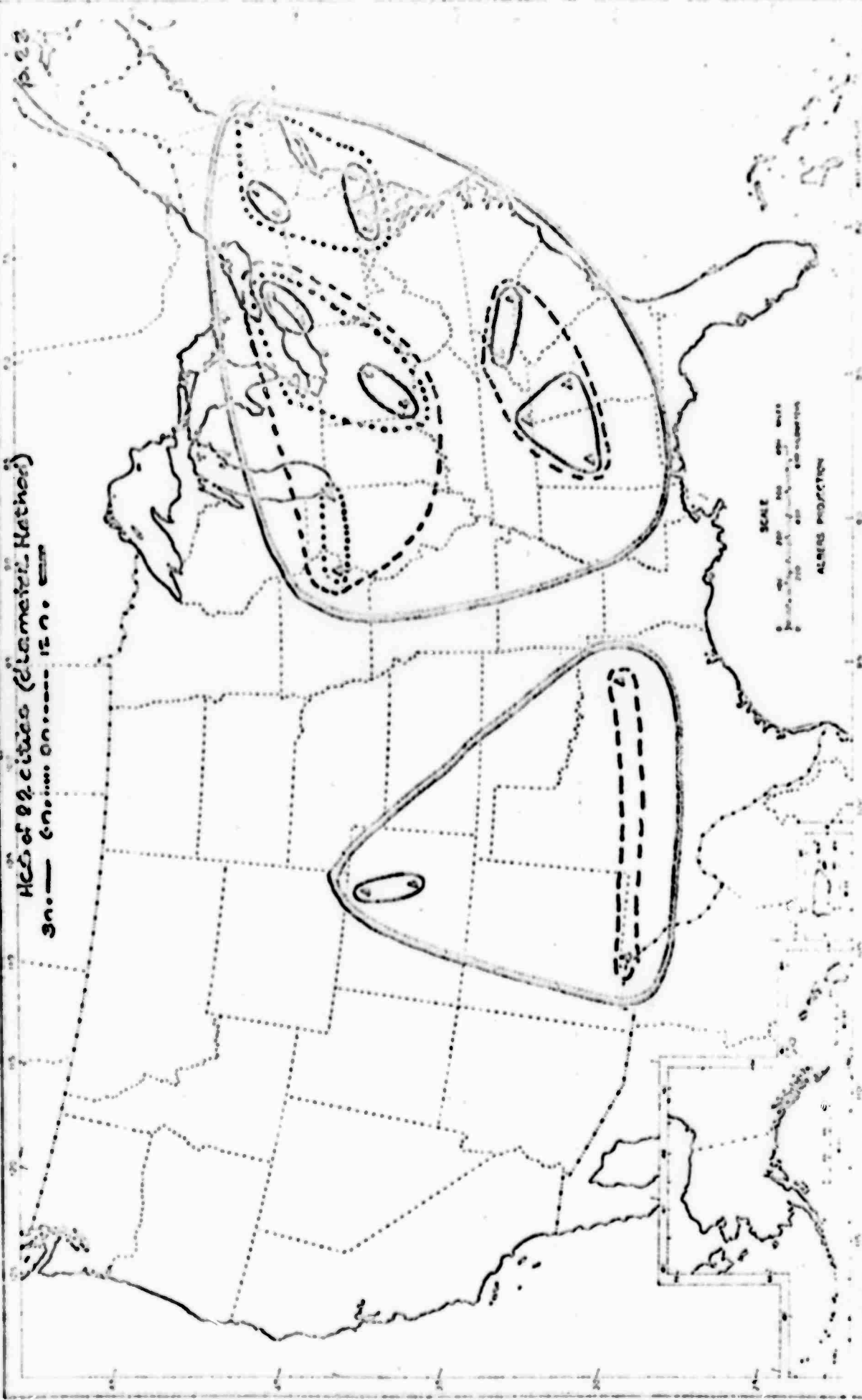


Figure VI. Dendrogram for 22 Cities

Diameter Method



HCS of 0.2 cities (diameters: Hatched)
 30. — 60. — 90. — 120. — 150. —

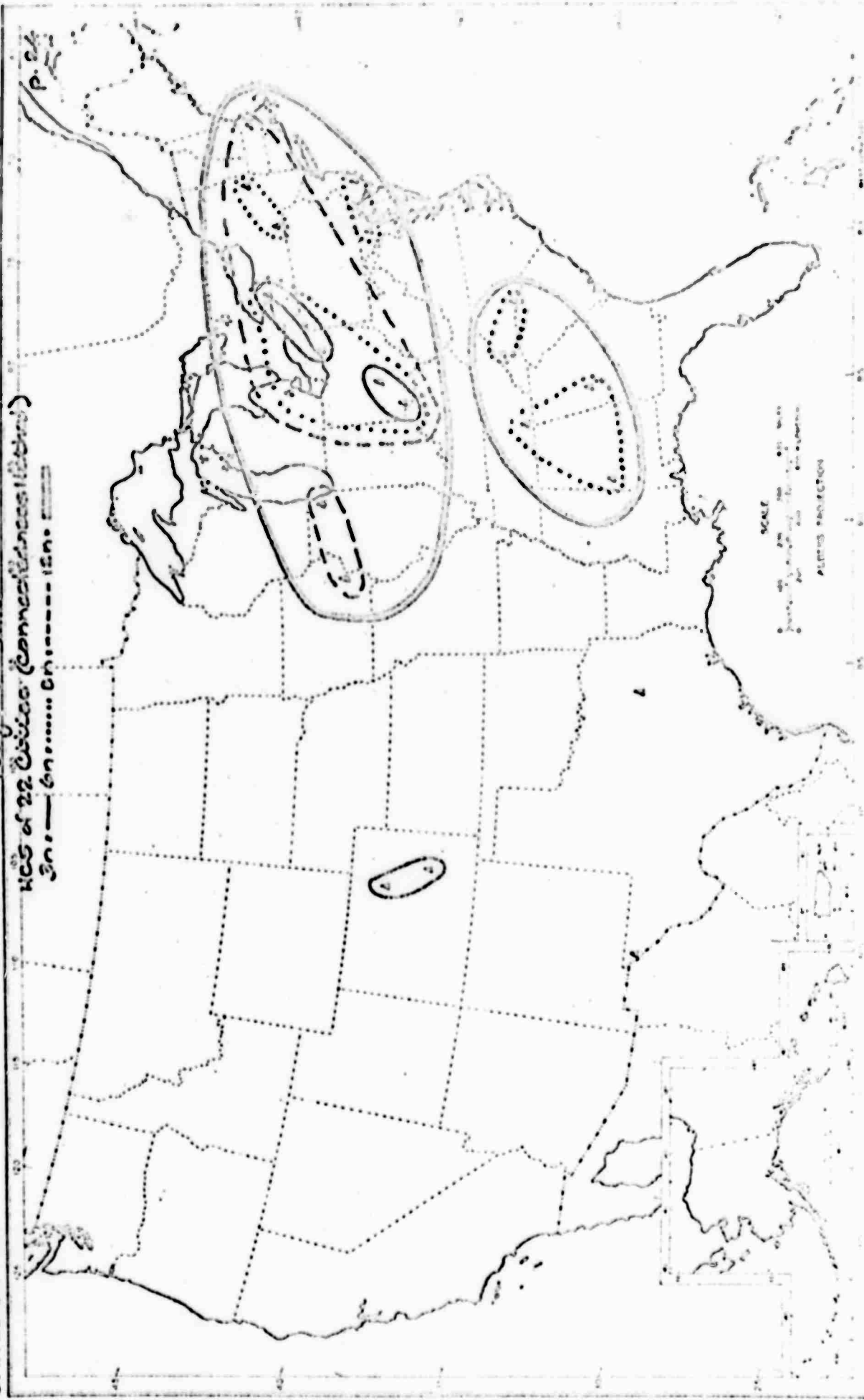


UNITED STATES

Figure VIII

No. 112

Map of 22 Counties (Connecticut, Illinois, Indiana, Michigan, Minnesota, Missouri, Nebraska, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Virginia, West Virginia, Wisconsin, Wyoming)



COVERED BY THE STATE OF NEW YORK

Page 1 of 1

but the subgroupings inside this area are not clear. For instance, Cincinnati and Boston are in the same group, but Atlantic City and Baltimore do not join this group until later. It would appear that in this case the diameter method groups cities into neater clusters and tends to account for all the cities at lower N levels than does the connectedness method.

The 60 city matrix of distances was analyzed in the same manner. Figures IX-X present the dendrograms and figures XI-XII display the geometric interpretations of the 60 city plasmode.

In the connectedness method, 16 cities are grouped only at the 60N level in the dendrograms, while all cities are accounted for in the diameter method prior to the 60N levels. Looking at figures XI and XII--the results of groupings presented in geometric representations--the connectedness method does not group several of the cities. That is, several cities do not belong to any of the groups delineated by the diameter method. It does group the large bloc of midwestern cities which did not come together in the factor analyses of this matrix, however. The diameter method (Figure XII) groups all the cities, but it results in some which are difficult to interpret. For instance, the divisions on the Kansas - Missouri, border are difficult to substantiate. The breakdown of group in the midwest is quite misleading.

2.3 Conclusions from Plasmodes

We have presented two different plasmodes. The first--22 cities--plasmode contained clearly recognizable groups. The second plasmode on 60 cities did not contain a simple grouping. There was a high density area in the midwest and an almost equal dispersion of cities on the map. Several suggestions can be made on the advisability of using each of the four types of analysis on a matrix.

It would appear that direct factor analysis of a rescaled distance matrix is more applicable than the HCS techniques in both the simple structure and the complexly structured case. When simple structure exists in the data, oblique

Figure IX. Dendrogram for 60 Cities

Diameter Method

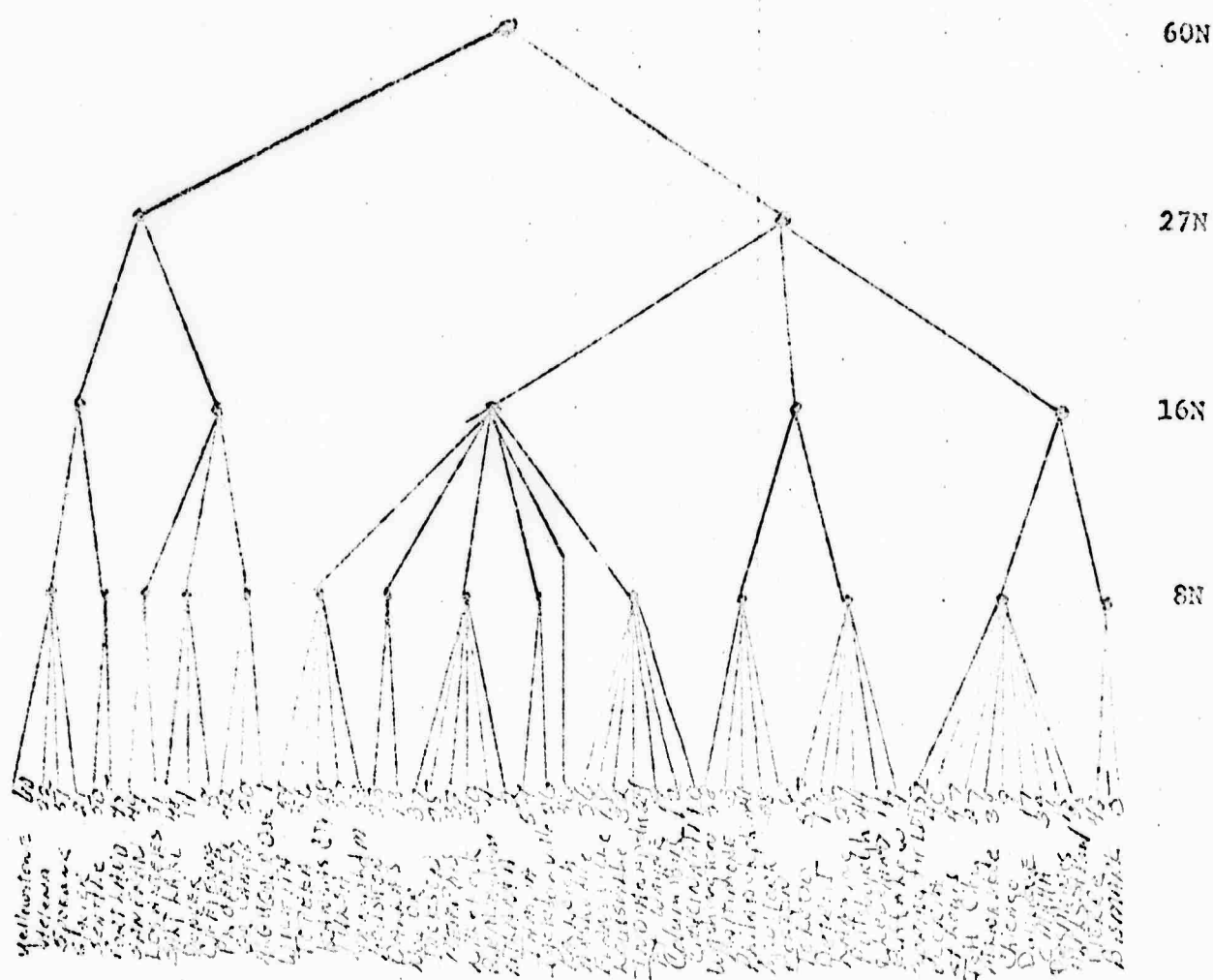
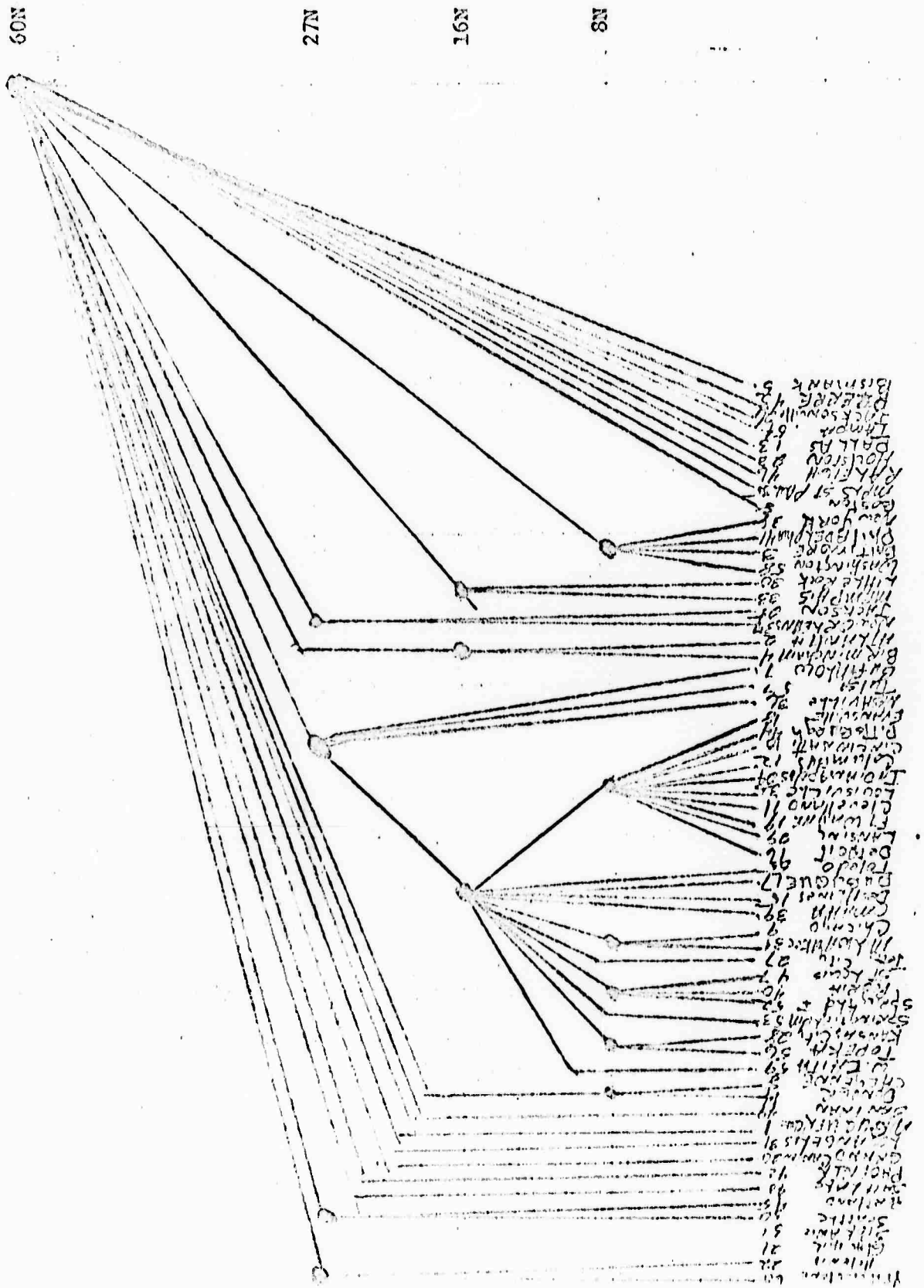
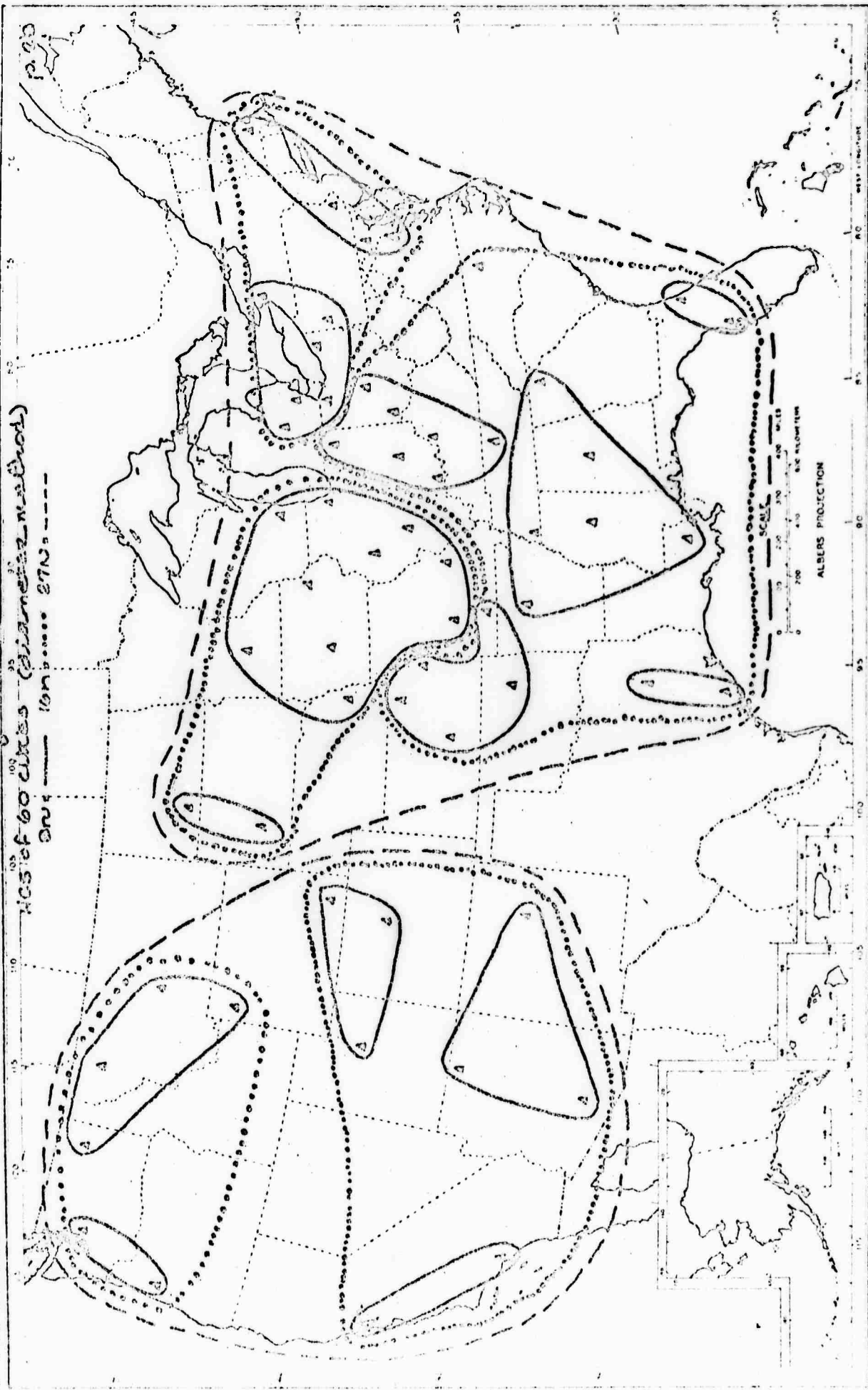


Figure X. Dendogram for 60 Cities
Connectedness Method

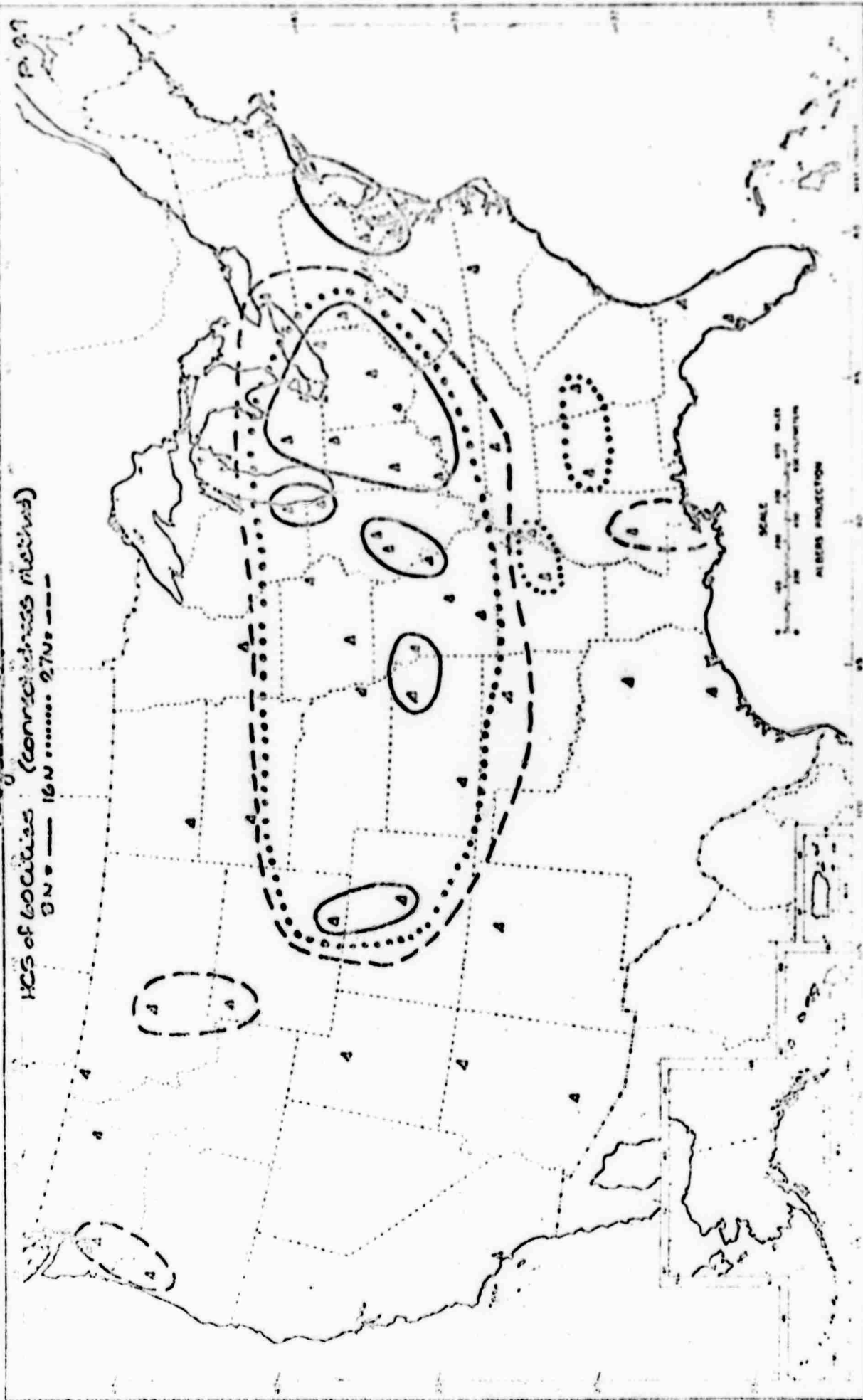




Figures VII

HCS of Locusts (Corroboration Method)

DN — 16N 27N —



dimensions are a better referent system than the orthogonal dimensions. Note the shrinkage of clusters from orthogonal to oblique in the 22 city plasmode. When the data points do not clearly cluster in space, orthogonal rotation appears more useful in defining groups than oblique factors. Oblique rotation is not employed by all analysts who use factor analysis. For example, Burt (1940, p. 266) has argued for orthogonal rotation rather than oblique. From our 2 plasmodes, it would seem that the argument is well founded when simple structure does not exist in the data.³

In the HCS analyses, it would appear that the diameter method is better when clusters are clear and unambiguous. The 22 city case points to this conclusion. On the other hand, the diameter method tends to break down when there is no apparent structure as in the 60 city plasmode. One distinct advantage of the HCS technique is that both criteria group cities in the highly dense midwestern section of the 60 city plasmode. The researcher who finds he has a large group of entities all clustering at one point in the space defined by his data matrix might wish to look at both connectedness and diameter methods and to choose the more interpretable of the two techniques, if he wishes to make distinctions within his highly clustered group.

To reiterate, the direct factor analysis techniques seem more appealing than HCS techniques in describing the structure of the space defined by the plasmodes. The choice between oblique or orthogonal solutions depends upon the complexity in the data. HCS is a useful technique when the researcher wishes to break up a dense cluster of entities with the diameter method being slightly more applicable in this case.

3. SUBSTANTIVE EXAMPLE

An analysis of conflict data will be presented to extend the findings of the plasmode experiments to international relations. Data have been collected on dyadic conflict behavior from the New York Times for 1963. The resulting matrix was 26 variables for dyadic conflict over 275 dyads. This matrix was factored (principal components) and rotated through orthogonal and oblique solutions (Hall and Rumel, 1965). Of the 275 nations in the study, only 61 had a significant amount of conflict on at least one of the five dimensions. Significance is defined as a factor score of 1.50 or higher. Two of the Hall and Rumel factors will be employed in this analysis: Unofficial incidence of violence and negative communication. Negative communication is indexed by such variables as accusations, and protests, while unofficial violence is related to attacks on embassies, persons, or the flag. Only these two dimensions were used in order to keep the geometric interpretations simple.

Factor scores for each of the 61 nations exhibiting significant conflict in the original matrix were organized into a new matrix. The order of this new matrix was 61 by 2: 61 nation dyads as rows and 2 factor columns. These scores locate each dyad in the space of the two conflict dimensions.

Once the dyads are located, the next step is to delineate their distance from each of the other dyads in the conflict space. The Euclidean distance measure has gained a good deal of support as a similarity measure (Cronbach and Gleser, 1953; Kennally, 1962). It measures both elevation (profile average) and scatter (profile standard deviation) similarity as well as similarity in profile shape. Thus, it determines precisely the congruence of spacial locations. The distance measure is:

$$d = \sum_{t=1}^p \sqrt{(S_{tB} - S_{tA})^2}$$

where S_{tB} = dyad B's score on the t^{th} factor

S_{tA} = Dyad A's score

p = the number of factor defining the space

In this case only two dimensions are employed. The technique is generalizable to as many dimensions as are required to define the space, however.

Once a matrix of distances has been developed, the clustering of dyads can be determined in the same manner as has been done in the plasmodes. Figure XIII shows the steps undergone in the analysis. The geometric interpretations of the orthogonal solution are presented in Figure XIV.

The orthogonal factors define three clusters. The major factor is centered around the axis. Much of the group is located in the negative quadrant of both negative communications and unofficial incidences of violence, however. This counter intuitive observation does not mean that the dyads display a negative amount of behavior used to name these factors. It signifies that while they have no occurrence of these forms of conflict, they do exhibit conflict behavior on other variables which have some slight negative loading on the two dimensions discussed here. The other two factors define groups of dyads which exhibit high amounts of either negative communication or unofficial acts of violence, but not both.

The information presented in the factor tables is also conveyed using a geometric interpretation. The coordinates in this presentation are the dimensions of dyadic conflict behavior delineated in the earlier study (Hall and Rummel, 1968). The dyads have been plotted in this space employing their factor scores

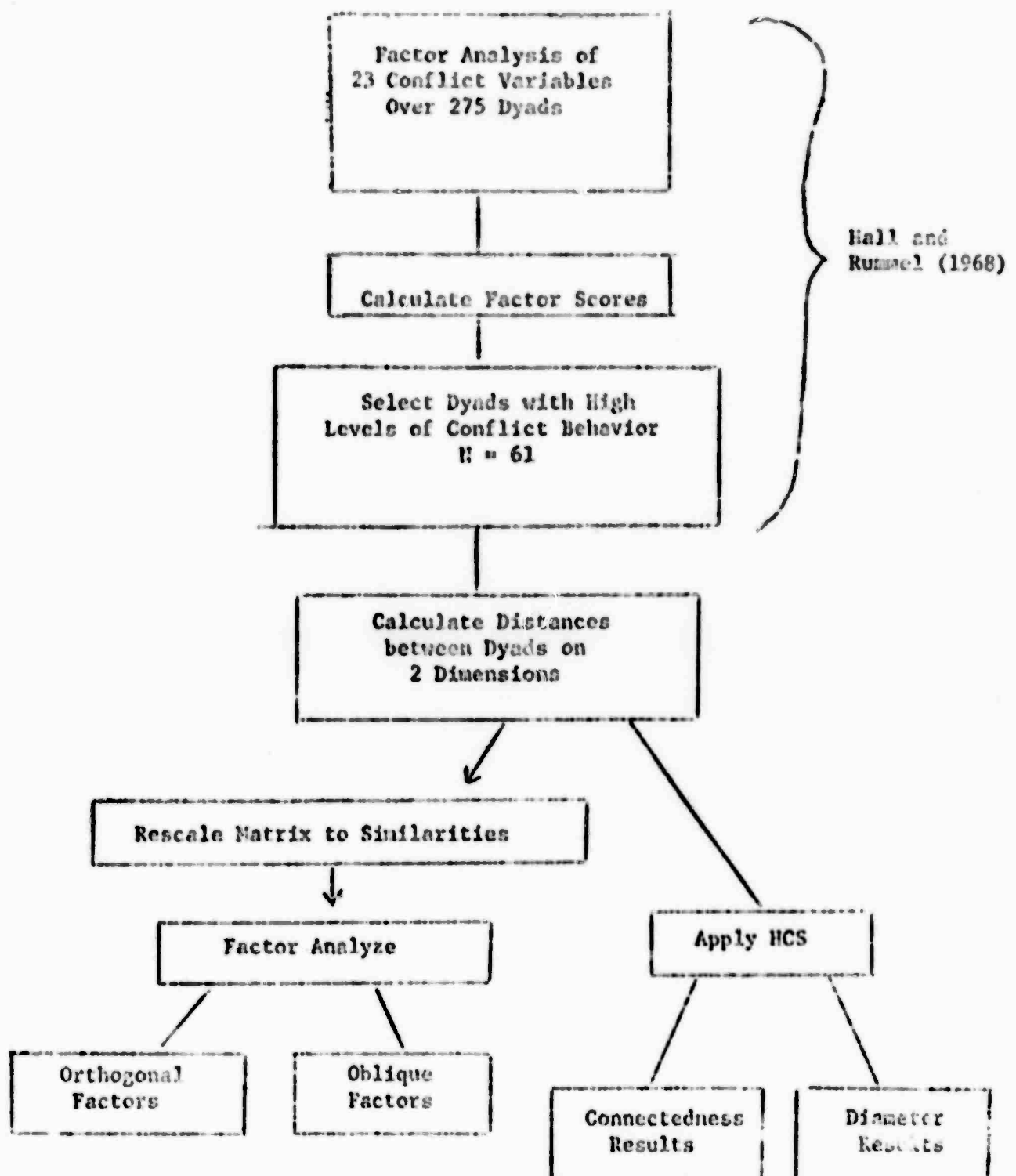
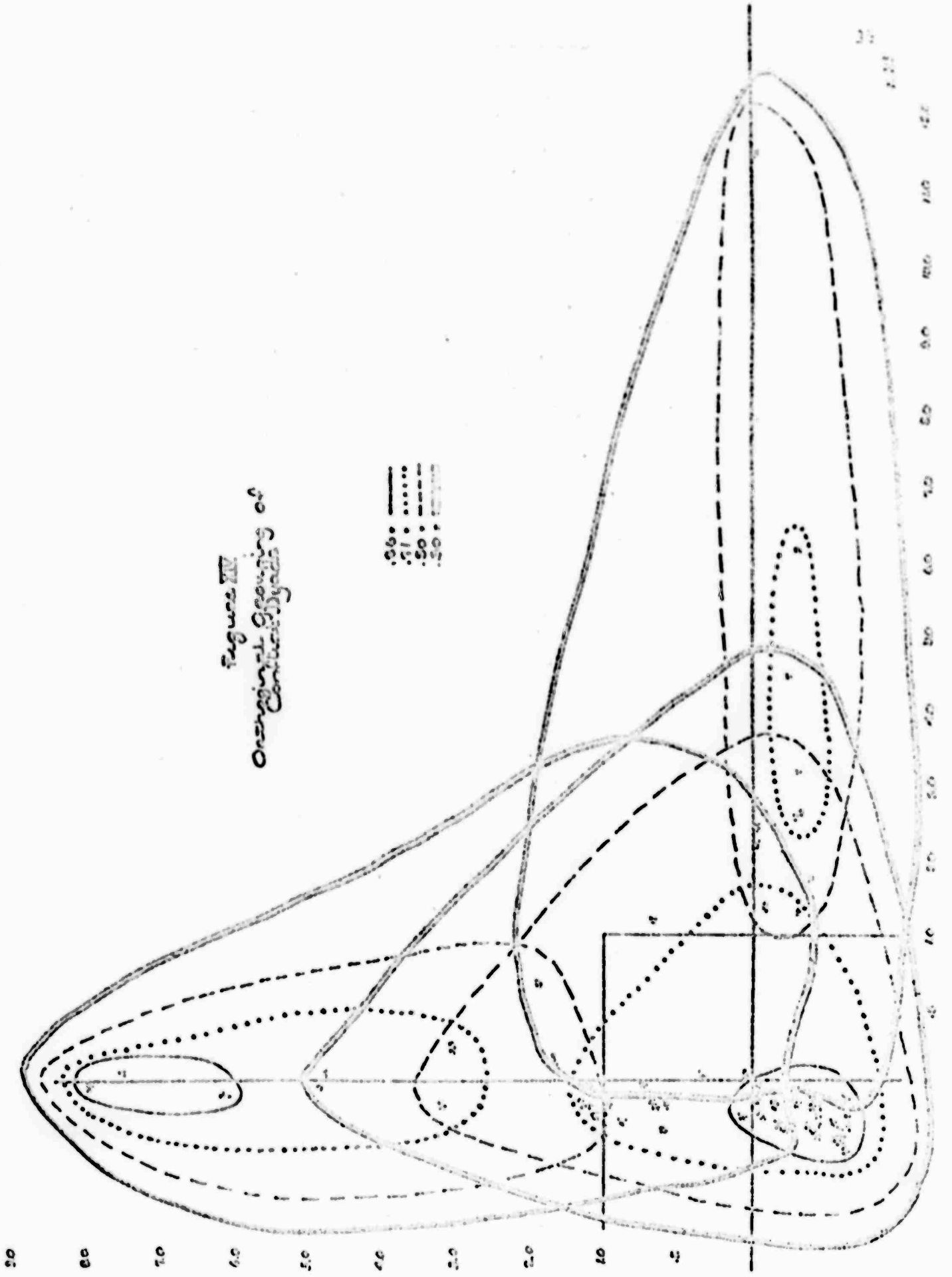


Figure XI'I

Table 6. Orthogonal Rotation of Conflict Dyads

Factor				Factor					
Dyads	1	2	3	Dyads	1	2	3		
1	USR-USA	1557	8902	0815	32	MAI-PHI	9112	2548	2619
2	CHN-USA	2526	8911	1332	33	EGP-ISR	8150	4444	2872
3	CHN-USR	3568	8561	1950	34	BRA-FRN	9015	3044	2617
4	IND-CHN	3543	8543	2073	35	JAP-USA	9095	2496	2755
5	FRN-USR	7293	5487	2929	36	IRQ-ISR	9128	2740	2637
6	UNK-USR	7327	5466	2889	37	USA-DOM	8994	2512	3018
7	ETH-SOM	7402	5370	2901	38	SEN-POR	9071	2488	2828
8	VEN-USA	-0423	-0456	7027	39	PAK-IND	7751	4962	2916
9	INS-UNK	2721	1754	3250	40	CAM-USA	8326	3932	3214
10	COL-USA	3993	2428	8114	41	ETH-FRN	9129	2676	2772
11	VEN-UNK	5079	2696	7682	42	USA-CUB	7899	4816	2800
12	IRN-USA	5651	2843	7310	43	BEL-FRN	8947	3088	2859
13	ECU-USA	5652	2850	7308	44	MOR-EGP	9047	2735	2928
14	CHT-JAP	5833	3168	7014	45	IND-UNS	9115	2550	2637
15	BEL-USA	6681	2920	6186	46	COF-USR	8932	2811	3093
16	YEM-UNK	7196	3008	5515	47	GUA-UNK	9118	2561	2731
17	VTS-USA	6865	4288	4999	48	VEN-HAI	9101	2572	2579
18	ALB-YUG	9125	2757	2652	49	ISR-SYR	7429	5396	3008
19	YUG-ALB	9114	2534	2732	50	CUB-USA	1359	8305	0624
20	CHN-CHT	9095	2678	2834	51	DOM-HAI	6180	5989	3846
21	CHT-CHN	9114	2658	2788	52	USR-CHN	5157	7434	2528
22	KOR-USA	7993	4579	3032	53	USA-ISR	5205	7255	2949
23	EGP-SAU	8775	3319	3001	54	IND-PAK	7540	5237	2874
24	ISR-JOR	8817	3286	2945	55	HAI-USA	7256	3282	5705
25	JOR-ISR	8757	3518	2742	56	LEB-SYR	8080	4513	2935
26	UNK-YEM	9030	2859	2884	57	UNK-SOM	9068	2739	2873
27	VTS-CAM	9030	2859	2884	58	USA-HAI	8846	3336	2789
28	SYR-ISR	8195	4374	2877	59	UNK-INS	8573	3474	3238
29	AL-INS	8233	4381	2691	60	USA-VTS	6802	5841	3190
30	FRN-BRA	9117	2779	2692	61	INS-MAI	5685	3165	6968
31	SOM-UNK	9033	3011	2680	62	PTAGE	9112	2793	2619



on both dimensions. The concentric circles or contours represent the factor loading criteria employed previously of .86, .71, .50, .30, respectively.

Turning to the oblique structure, we are able to interpret the primary pattern matrix. The clustering of the two groups representing extreme behavior on either of the dimensions is more specific to the extreme ends of the conflict factors. This is the case in the orthogonal presentation. The group clustering around the axis has enlarged slightly. There seems to be a tendency toward one factor when a large percentage of the points are grouped together. In contrast to the 60 city plasmode there are two definite clusters which remain equally important in the oblique rotation as they were originally in the orthogonal case.

The angle between the two conflict behavior dimensions is 90 degrees in the geometric presentation. This reflects the fact that the dimensions were actually independent (Hall and Rumel, 1968).

A factor correlation matrix is computed in oblique rotation that delineates the amount of correlation between the factors. Hall and Rumel found a 0.00 correlation between negative communication and unofficial incidence of violence. The same procedure was employed to ascertain the relationships between the group factors in this analysis. If the factors are highly inter-correlated, each group is overlapping the other groups and the factors are not equally defining a cluster of dyads.

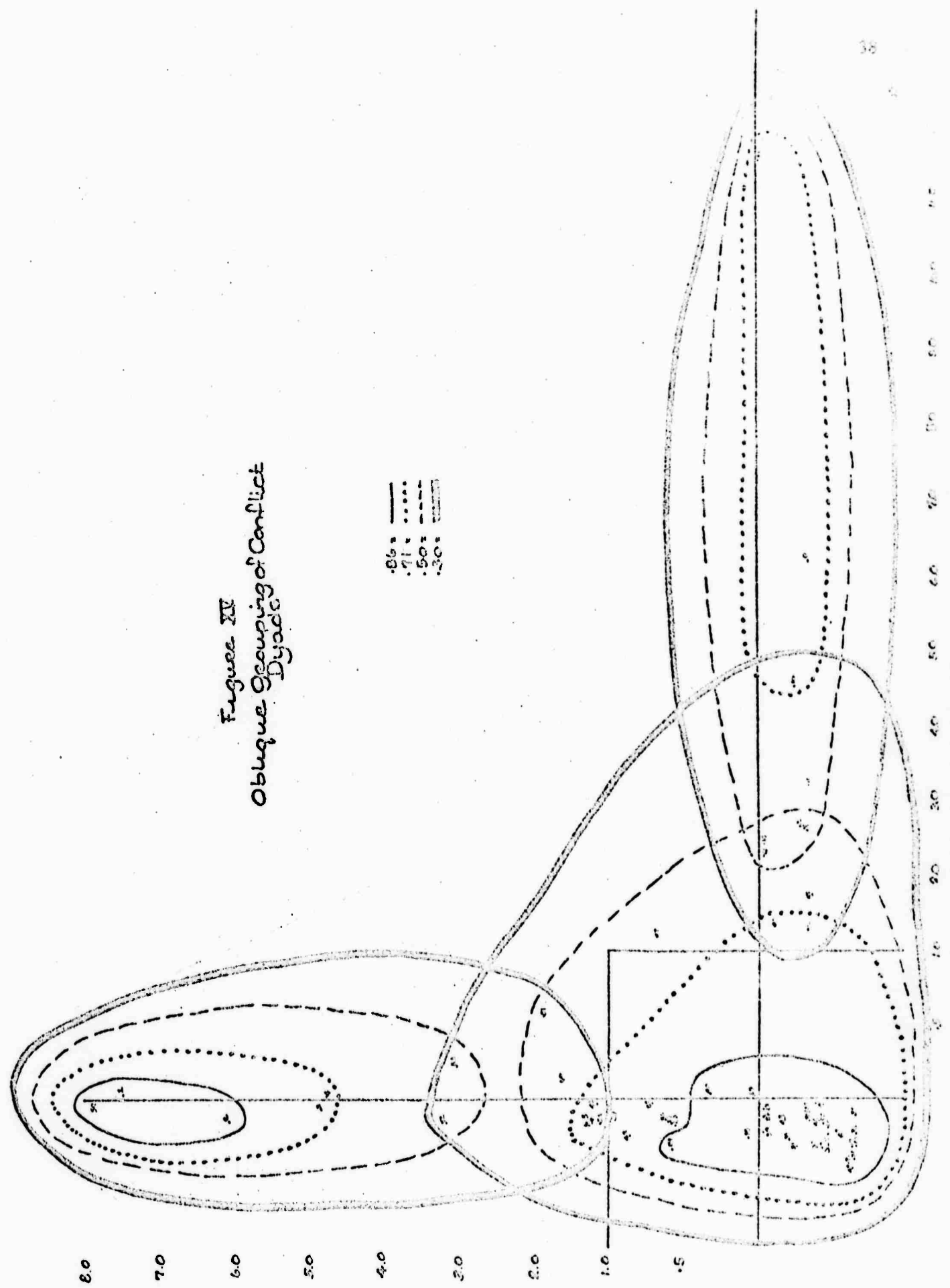
Correlations Between Primary Factors:

Factor	1	2	3
1	-----		
2	.475	-----	
3	.422	.255	-----

Table 7. Oblique Rotation of Conflict Dyads

Dyads		Factor			Dyads		Factor		
		1	2	3			1	2	3
1	USR-USA	0081	9127	-0388	32	MAL-PHI	1.0024	-0388	-0061
2	CHN-USA	1162	8783	-0118	33	EGP-ISR	8483	1972	0303
3	CHN-USR	2384	8016	0270	34	BRA-FRN	9806	0179	-0005
4	IND-CHN	2340	7995	0410	35	JAP-USA	9995	-0448	0083
5	FRN-USR	7264	3382	0516	36	IRQ-ISR	1.0006	-0184	-0074
6	UNK-USR	7314	3351	0466	37	USA-DOM	9835	-0418	0390
7	ETH-SOM	7419	3222	0466	38	SEN-POR	9957	-0455	0179
8	VEN-USA	-1402	-0891	7637	39	PAK-IND	7910	2663	0418
9	INS-UNK	1764	0401	7831	40	CAN-USA	8746	1321	0663
10	COL-USA	3174	0732	7259	41	NTH-FRN	1.0007	-0261	0022
11	VEN-UNK	4480	0703	6462	42	USA-CUB	8130	2464	0265
12	IRN-USA	5187	0705	5889	43	BEL-FRN	9694	0237	0182
13	ECU-USA	5187	0713	5886	44	MOR-EGP	9871	-0185	0248
14	CHT-JAP	5384	1022	5492	45	IND-UNS	1.0026	-0387	-0053
15	YOL-USA	6558	0539	4389	46	COP-USR	9695	-0078	0461
16	YEM-UNK	7249	0518	3519	47	GUA-UNK	1.0011	0385	0066
17	VTS-USA	6688	2059	2950	48	VEN-HAI	1.0013	-0354	-0113
18	ALB-YUG	9997	-0166	-0059	49	ISR-SYR	7449	3135	0578
19	YUG-ALB	1.0014	-0411	0049	50	CUB-USA	-0111	9100	-0525
20	CHN-CHT	9950	-0256	0150	51	DOM-HAI	5714	4222	1767
21	CHT-CHN	9983	-0280	0098	52	USR-CHN	4409	6228	0529
22	KOR-USA	8254	2158	0506	53	USA-USR	4441	5985	0979
23	EGP-SAU	9426	0534	0362	54	IND-PAK	7613	3034	0409
24	ISR-JOR	9490	0488	0293	55	HAI-USA	7299	0813	3254
25	JOR-ISR	9403	0776	0074	56	LEB-SYR	8383	2065	0384
26	UNK-YEM	9832	-0041	0206	57	UNK-SOM	9901	-0184	0194
27	VTS-CAN	9832	-0041	0206	58	USA-HAI	9537	0545	0015
28	SAR-ISR	8554	1881	0302	59	UNK-INS	9123	0750	0658
29	MAL-INS	8624	1890	0092	60	USA-VTS	6575	3908	0903
30	FRN-BRA	9978	-0142	-0016	61	INS-MAL	5453	1006	5428
31	SOM-UNK	9836	0139	-0026	62	PEACE	9968	-0126	-0009

Figures III
Oblique Grouping of Conflict



As the geometric interpretations would lead us to expect, the two extreme factors are more closely related to the factor centered around the axis than they are to each other. Just a little over 6 per cent of the variance of the two extreme factors are in common while better than 16 per cent of the variance in both cases is in common with the factor centered around the axis.

Another test of the usefulness of this technique is to designate the dyads which exhibited high amounts of behavior on either of the conflict dimensions. Dyads displaying a high amount of negative behavior were: Cuba to United States, Soviet Union to United States; China to United States; China to Soviet Union; and Indonesia to China. The high incidence of negative communications in these directed dyads is certainly plausible. Turning to high levels of unofficial violence, these dyads are: Venezuela to United States; Indonesia to United Kingdom; Columbia to United States and Venezuela to United Kingdom. Venezuelan rebels (FLAN) staged several anti-American and British attacks in early 1963. Indonesians also demonstrated against the British sponsored federation of Malaysia.

The hierarchical clustering technique was also applied to the distance matrix derived prior to factor analysis. Both the dendrograms and geometric interpretation will be presented here. The connectedness technique is presented first. The countour lines for N levels are derived by obtaining the mean size of each of the factor contours (at loading levels of .86, .71, .50, .30). Quite a few of the dyads, especially those with high scores are not grouped using the connectedness technique. The groups themselves tend to be intuitively acceptable. There are exceptions, however. Dyads 42 and 39 (USA-Cuba; Pak-Ind) do not seem to be separate spacially from the other large group near them.

Diameter Case

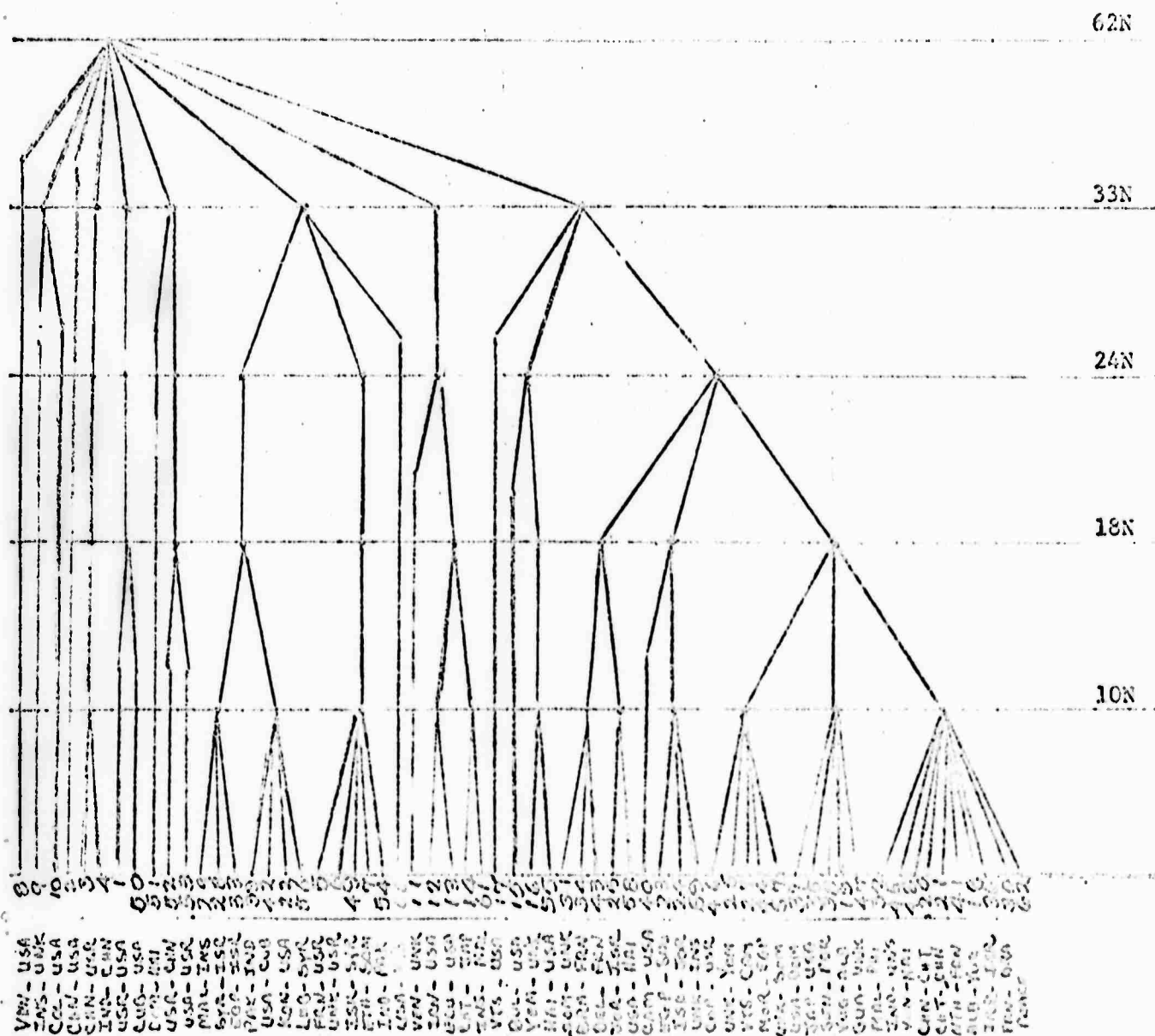


Figure XVII. Dendogram for Conflict Dyads
Connectedness Method

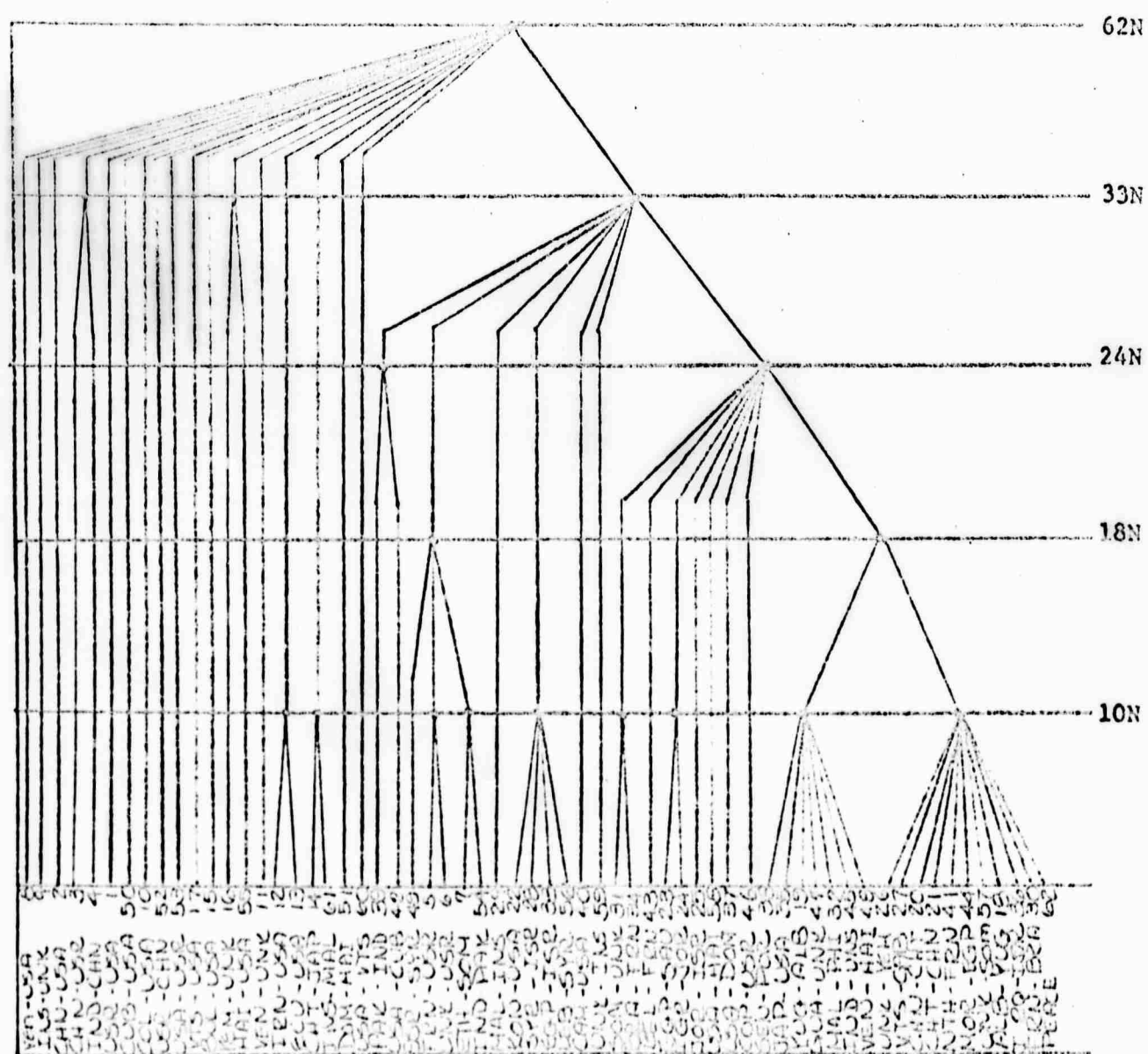


Figure XVII
HCS Grouping of Conflict Dyads
(diameter method)

10N = ———
18N =
24N = - - - -
33N = ———

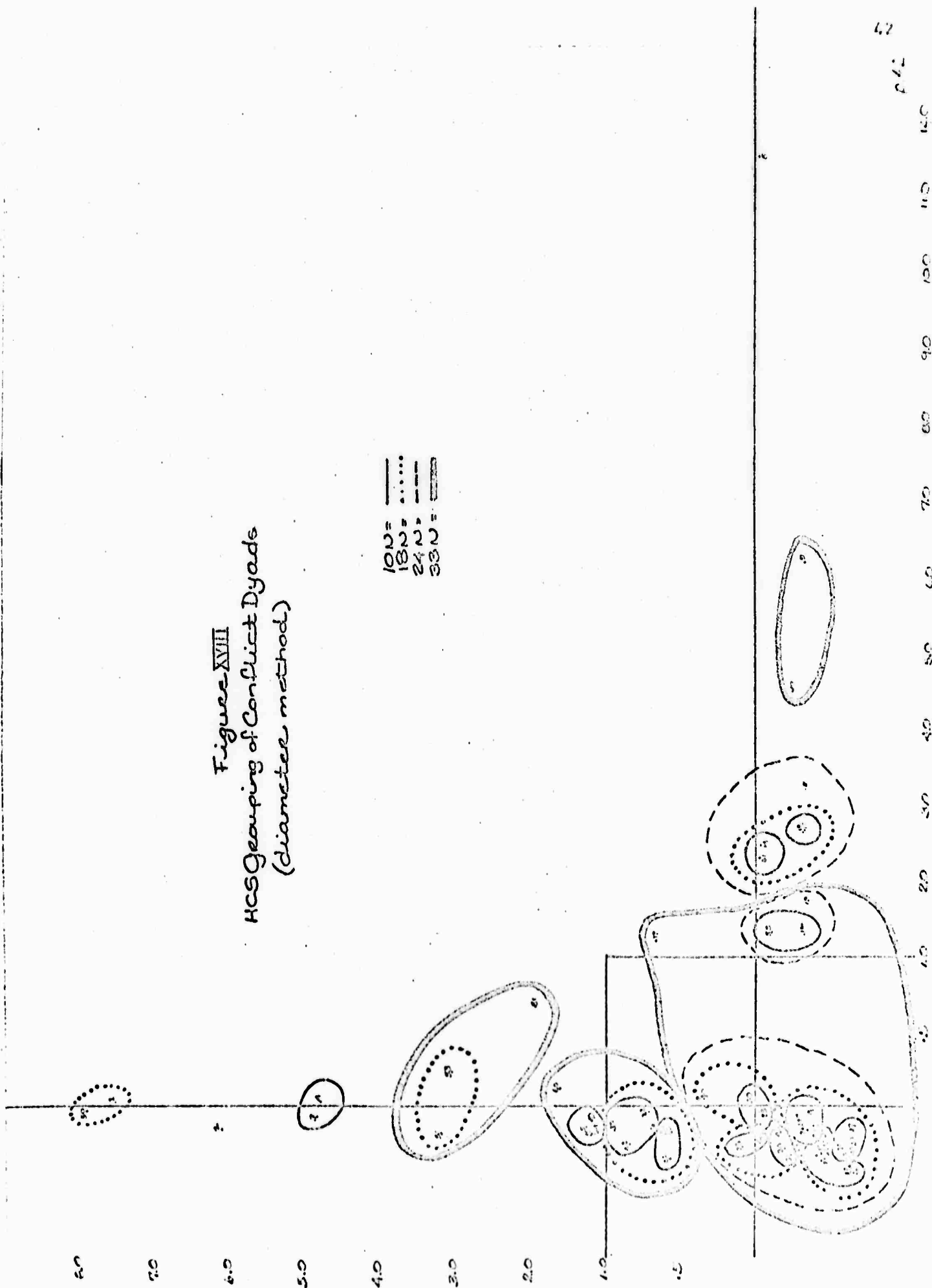
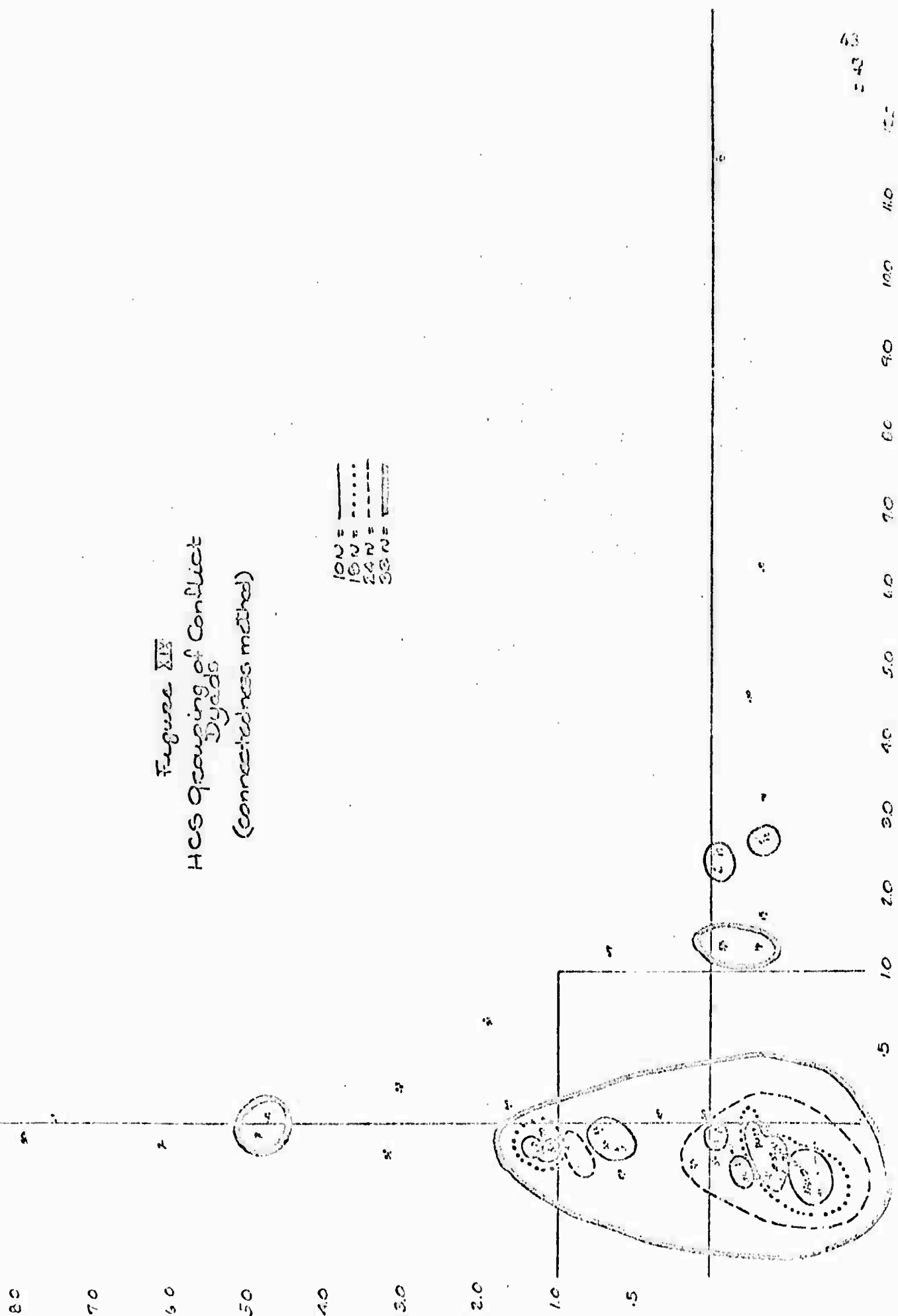


Figure XIX
HCS Grouping of Conflict
Dyads
(connectedness method)

10N = ———
18N =
24N = - - - -
32N = ———



The diameter method clusters almost all dyads before the last level. The groups seem more interpretable and would suggest that in this case the diameter method is more applicable than the connectedness method. The clusters seem to define groups in terms of the incidence of either negative behavior or unofficial violence that the dyads exhibited.

This case seems to be one in which there is a clearly defined structure to the matrix. Therefore, the groups are quite distinct. There is one general group with a series of smaller groups on both of the original conflict dimensions. We have a combination of the characteristics in the two plasmode. There are clearly recognizable groups in the conflict case as in the 22 city plasmode. On the other hand, the conflict case displays a high density area similar to the mid-western group of cities in the 60 city plasmode. Clearly both orthogonal and oblique factors are interpretable in this case. The choice depends upon the substantive interests of the analyst. Factor analysis has grouped the dyads with little activity on these two dimensions into a large cluster. Hierarchical clustering makes finer distinction within this group.

In HCS analysis it would appear that the diameter method has produced clear and unambiguous cluster. The connectedness method has left many dyads unaccounted for and has produced some counter intuitive groupings.

These findings are consistent with those in the plasmode experiments. The researcher is advised to investigate the usefulness of both direct factor analysis and hierarchical clustering techniques to his analysis before using only one or the other. The ultimate test of the usefulness of either technique rests in the ability to organize or explain new data independently collected.

BIBLIOGRAPHY

1. Ball, G. F., 1965, "Data Analysis in the Social Sciences: What About the Details?" in Proceedings -- Fall Joint Computer Conference 1965.
2. Brans, S. J., 1966, "Transaction Flows in the International System," American Political Science Review, Vol. LX, (Dec. 1966) 4, pp. 880-899.
3. Burt, C. L., 1941, The Factors of the Mind, New York: MacMillan Co.
4. Cattell, R. B., and Joseph Jaspers, "A General Plasmode (No. 30-10-5-2) for Factor Analytic Exercises and Research," Multivariate Behavioral Research Monograph No. 67-3.
5. _____, 1966, "The Scree Test for the Number of Factors," Multivariate Behavioral Research, 1966, 1, pp. 245-276.
6. Cronbach, L. J. and G. C. Gleser, 1963, "Assessing Similarity Between Profiles," Psychological Bulletin, 50, pp. 456-473.
7. Davis, P. J., 1965, The Mathematics of Matrices, Waltham, Mass.: Blaisdell Publishing Company.
8. Hall, Dennis, "Computer Program Profile," Research Report #14, Dimensionality of Nations Project, mimeo, 1968.
9. Hall, Dennis and R. J. Rummel, 1968, "The Patterns of Dyadic Foreign Conflict Behavior for 1963," The Dimensionality of Nations Project, Research Report #12, mimeo.
10. Harman, H. H., 1967, Modern Factor Analysis, Revised ed. Chicago: University of Chicago Press.
11. Horst, P., 1963, Matrix Algebra for Social Scientists, New York: Holt, Rinehart & Winston.
12. Johnson, S. C., 1967, "Hierarchical Clustering Schemes," Psychometrika, Vol. 32, #3 (September 1967), pp. 241-254.
13. Nunnally, J., 1962, "The Analysis of Profile Data," Psychological Bulletin, Vol. 59, No. 4., pp. 311-319.
14. Russett, Bruce, 1967, International Regions and the International System: A Study in Political Ecology, Chicago: Rand McNally.
15. Rummel, R. J., 1969, Applied Factor Analysis, Evanston, Northwestern University Press.
16. Sokal, Robert R., and Peter H. W. Sneath, 1963, Principles of Numerical Taxonomy, San Francisco: W. H. Freeman and Company.
17. Ward, J. H. Jr., 1963, "Hierarchical Grouping to Optimize an Objective Function," Journal of the American Statistical Association, 58, pp. 236-44.

Footnotes

1. See Geoffrey H. Ball, "Data Analysis in the Social Sciences: What about the Details?" in Proceedings -- Fall Joint Computer Conference, (1965).
2. For a description of this approach to grouping, see Rummel, Applied Factor Analysis, Chapter 22 . (1969).
3. For a comparison of orthogonal and oblique rotation and a discussion of the interpretability of each, see Rummel, 1969, Chapter 16 and Harman, H. (1967).
4. For details of matrix manipulation, see Davies (1965), or Horst, 1963.
5. Biquartimin solutions were calculated.